****

Exploratory Data Analysis

****

Size of the DataSet:

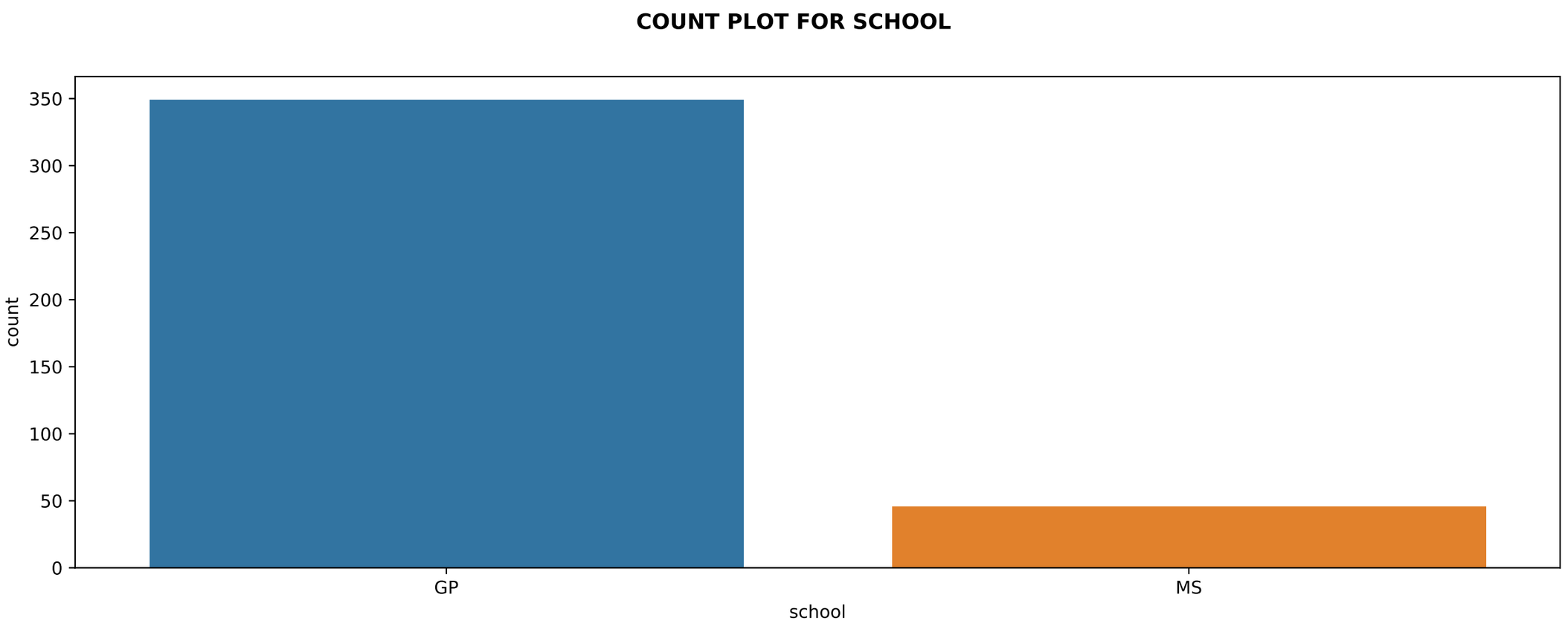
Attribute	Value
No of Rows	395
No of Columns	32
No of Duplicate Rows	0
Duplicate Columns	None

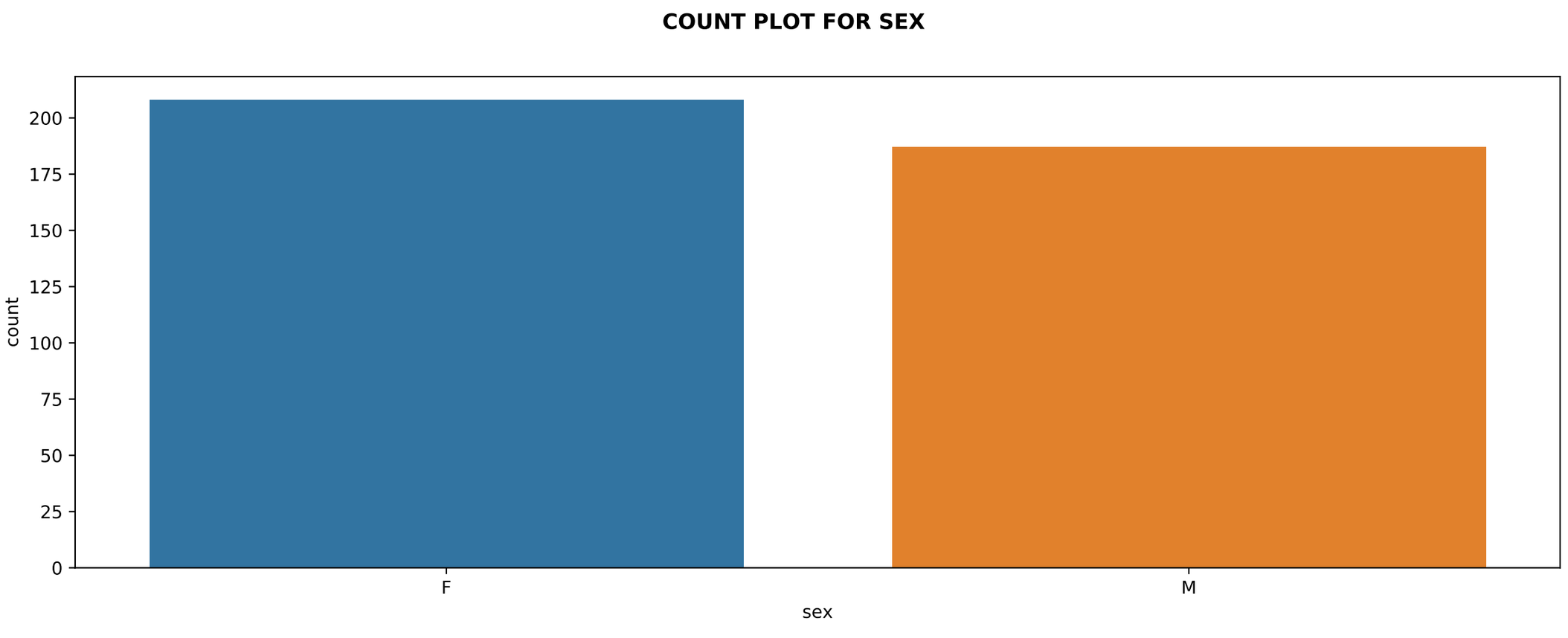
COLUMN DETAILS

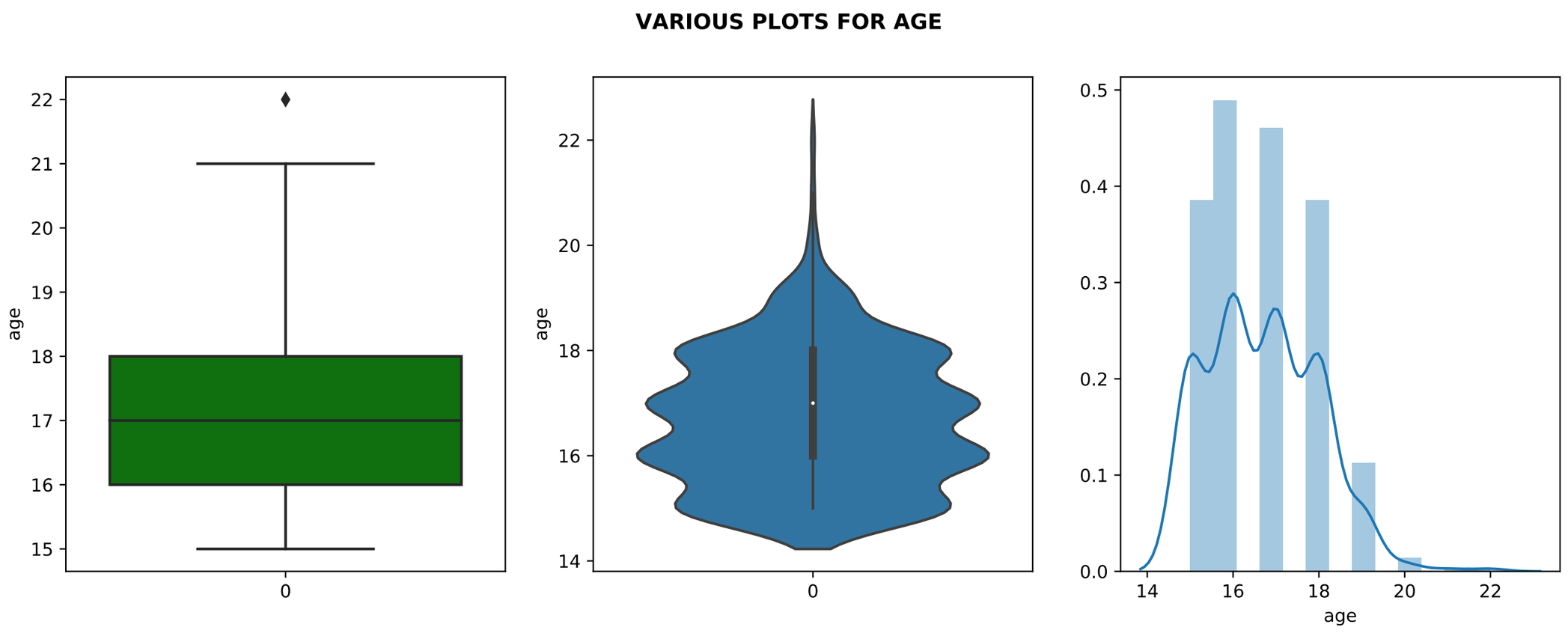
Column Name	Data Type	IQR	Outliers	Has Null	Median	Mode
age	int64	2.0	Outliers Present	No Null	17.0	16
Medu	int64	2.0	No Outliers	No Null	3.0	4
Fedu	int64	1.0	Outliers Present	No Null	2.0	2
traveltime	int64	1.0	Outliers Present	No Null	1.0	1
studytime	int64	1.0	Outliers Present	No Null	2.0	2
failures	int64	0.0	Outliers Present	No Null	0.0	0
famrel	int64	1.0	Outliers Present	No Null	4.0	4
freetime	int64	1.0	Outliers Present	No Null	3.0	3
goout	int64	2.0	No Outliers	No Null	3.0	3
Dalc	int64	1.0	Outliers Present	No Null	1.0	1
Walc	int64	2.0	No Outliers	No Null	2.0	1
health	int64	2.0	No Outliers	No Null	4.0	5
absences	int64	8.0	Outliers Present	No Null	4.0	0
G1	int64	5.0	No Outliers	No Null	11.0	10
G2	int64	4.0	Outliers Present	No Null	11.0	9

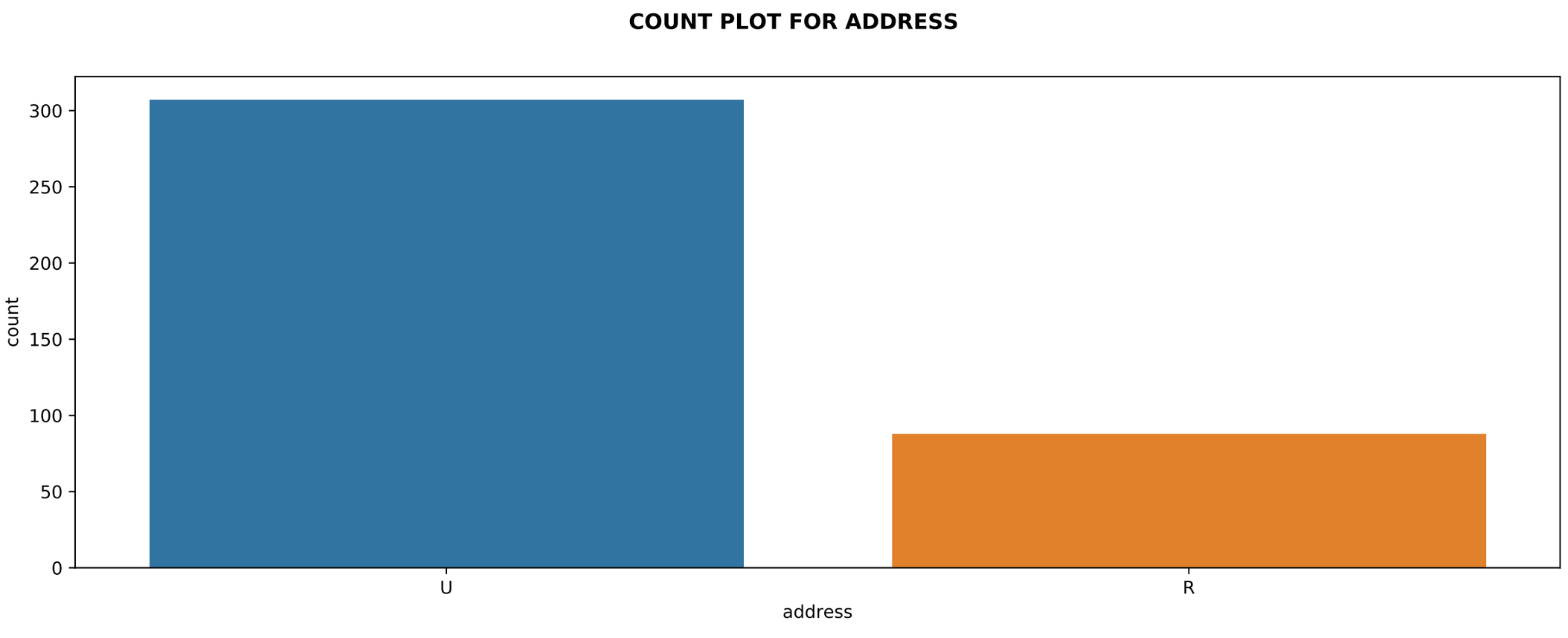
DATASET STATISTICS

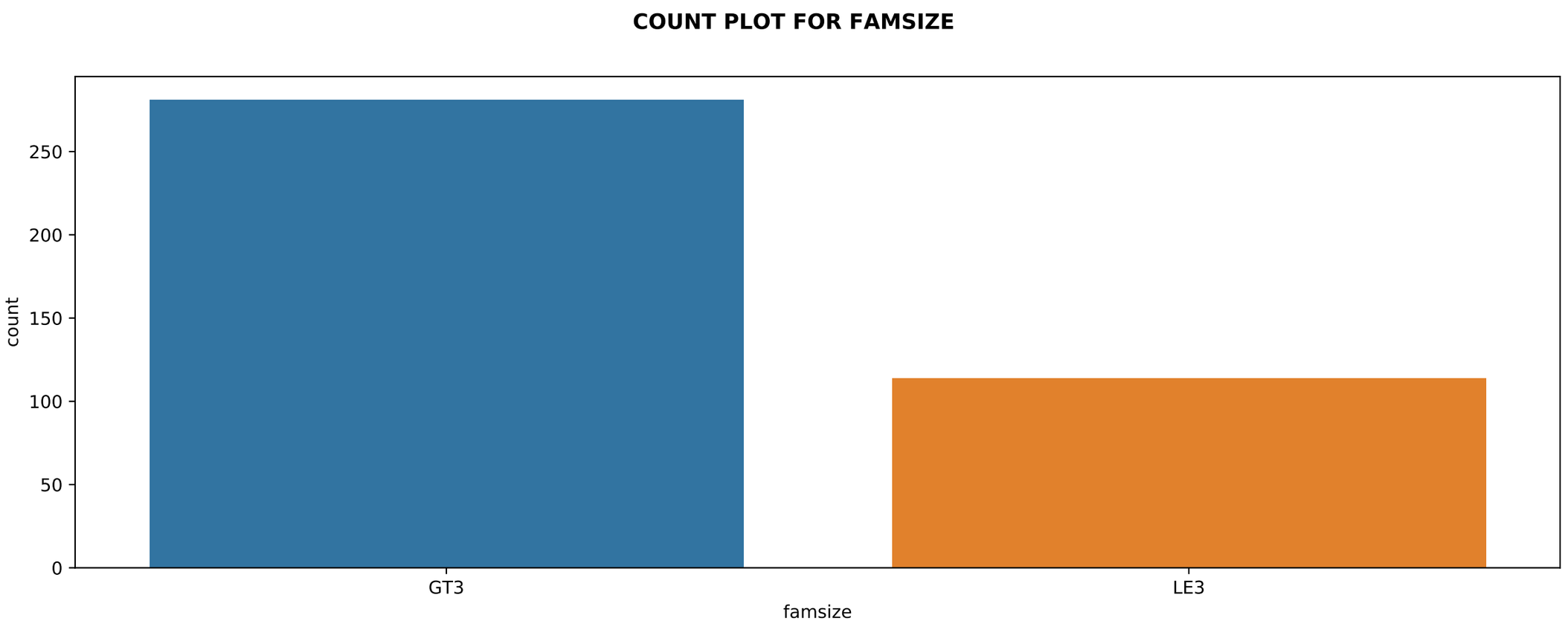
Column	count	mean	std	min	25%	50%	75%	max
age	395.0	16.696	1.276	15.0	16.0	17.0	18.0	22.0
Medu	395.0	2.749	1.095	0.0	2.0	3.0	4.0	4.0
Fedu	395.0	2.522	1.088	0.0	2.0	2.0	3.0	4.0
traveltime	395.0	1.448	0.698	1.0	1.0	1.0	2.0	4.0
studytime	395.0	2.035	0.839	1.0	1.0	2.0	2.0	4.0
failures	395.0	0.334	0.744	0.0	0.0	0.0	0.0	3.0
famrel	395.0	3.944	0.897	1.0	4.0	4.0	5.0	5.0
freetime	395.0	3.235	0.999	1.0	3.0	3.0	4.0	5.0
goout	395.0	3.109	1.113	1.0	2.0	3.0	4.0	5.0
Dalc	395.0	1.481	0.891	1.0	1.0	1.0	2.0	5.0
Walc	395.0	2.291	1.288	1.0	1.0	2.0	3.0	5.0
health	395.0	3.554	1.39	1.0	3.0	4.0	5.0	5.0
absences	395.0	5.709	8.003	0.0	0.0	4.0	8.0	75.0
G1	395.0	10.909	3.319	3.0	8.0	11.0	13.0	19.0
G2	395.0	10.714	3.762	0.0	9.0	11.0	13.0	19.0

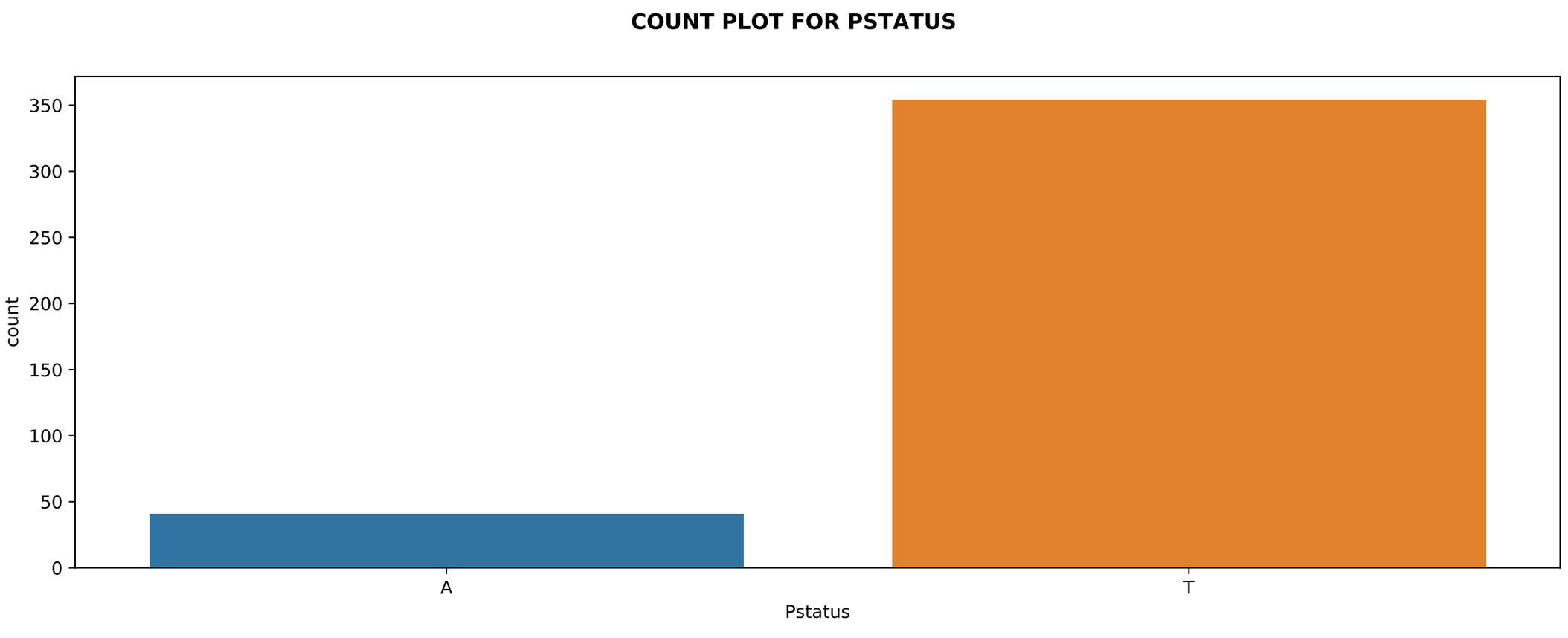

COUNT PLOT FOR SCHOOL
350
300
250
200
150
100
50
0
count
GP
MS
school

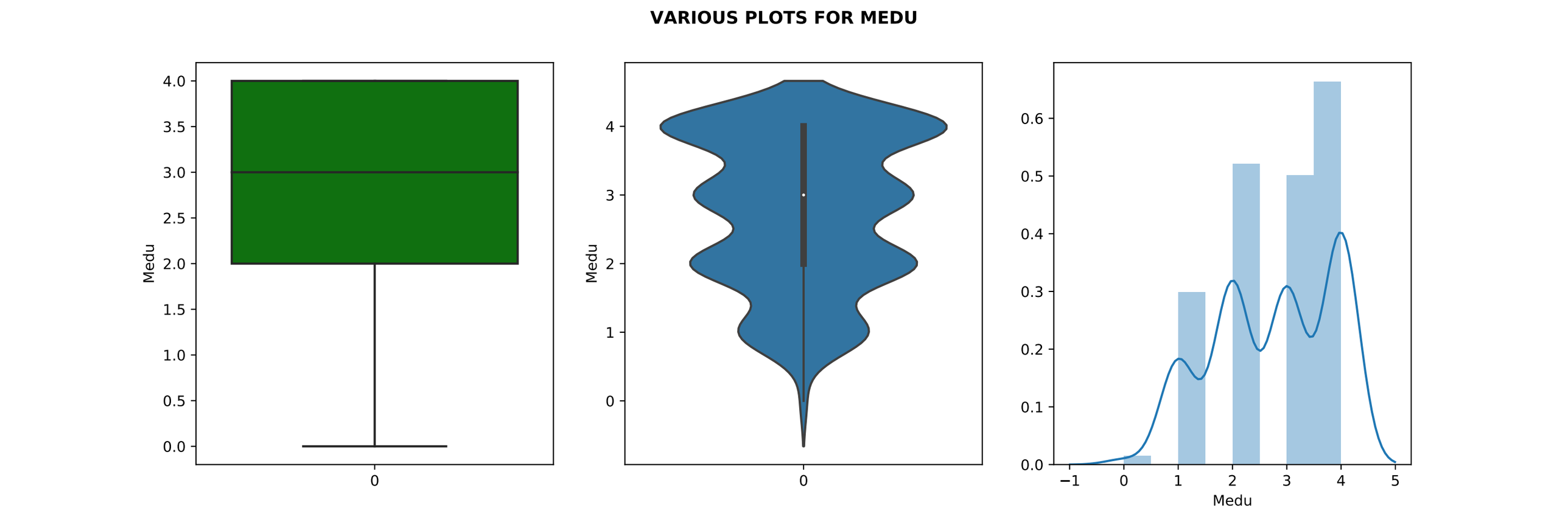

COUNT PLOT FOR SEX
count
0
25
50
75
100
125
150
175
200
F
M
sex

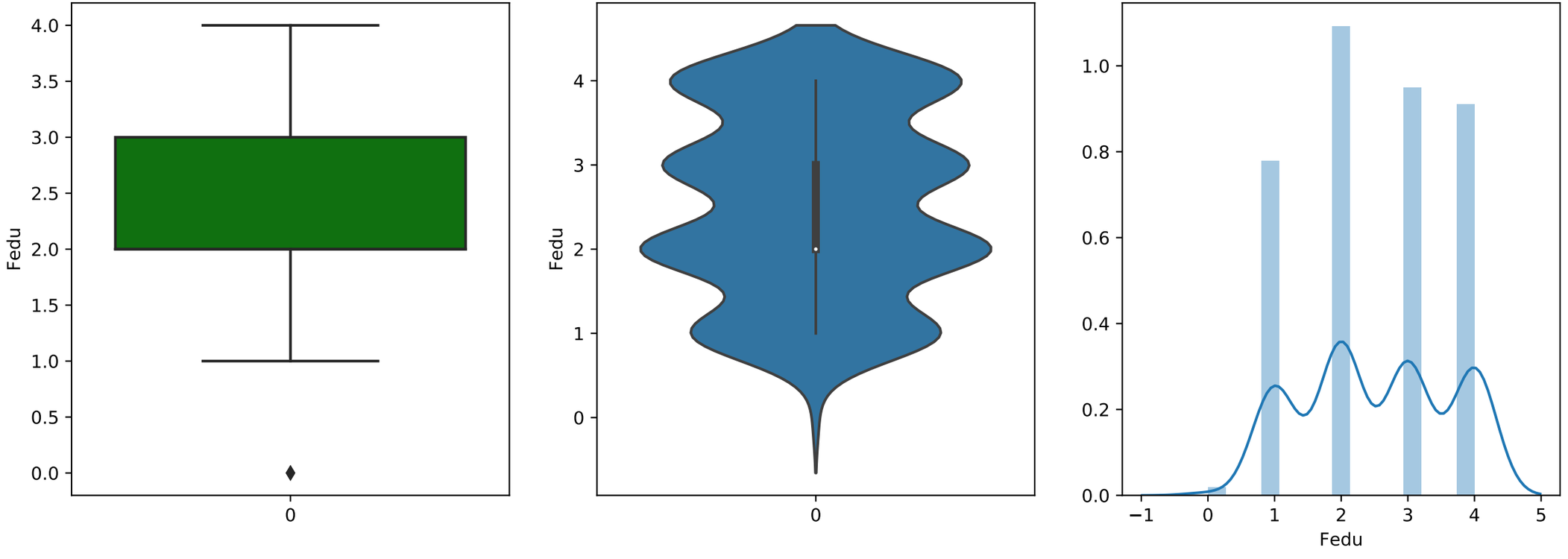

VARIOUS PLOTS FOR AGE
age
0
22
21
20
19
18
17
16
15
14
0.5
0.4
0.3
0.2
0.1
0.0

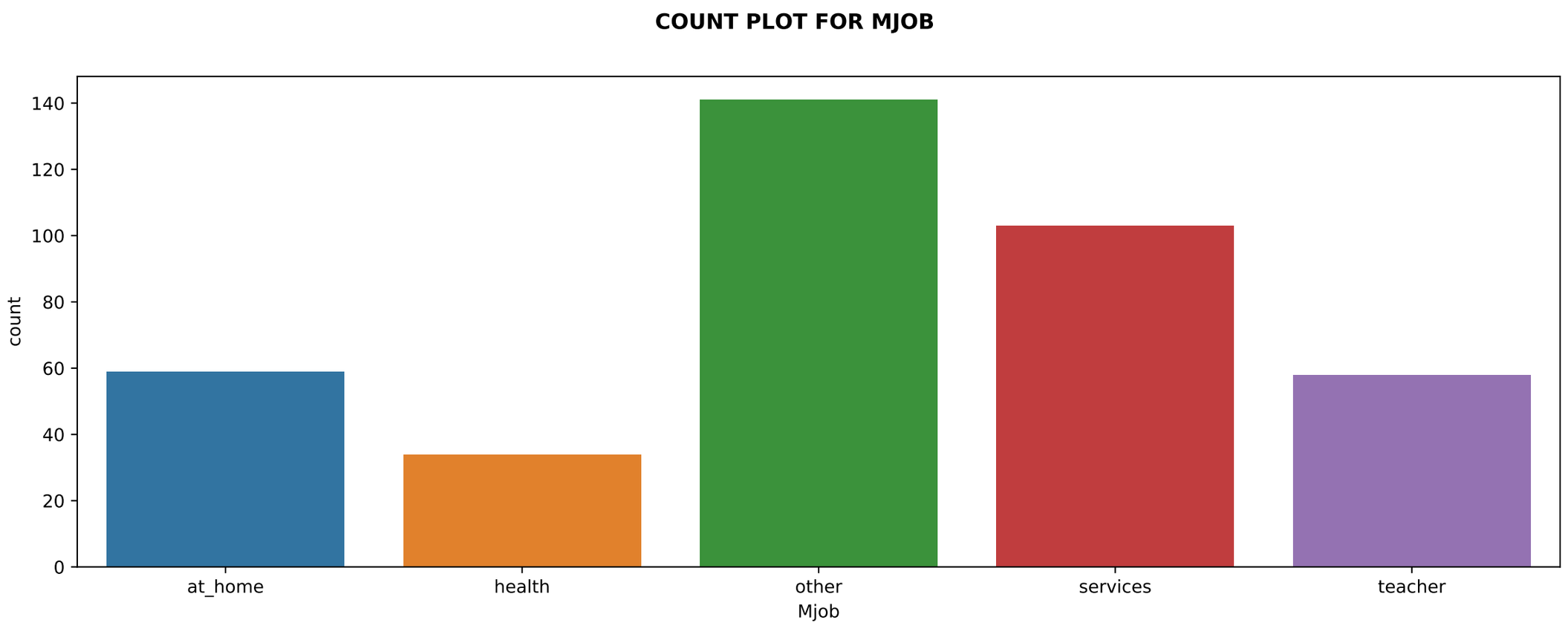

COUNT PLOT FOR ADDRESS
300
250
200
150
100
50
0
count
U
R
address

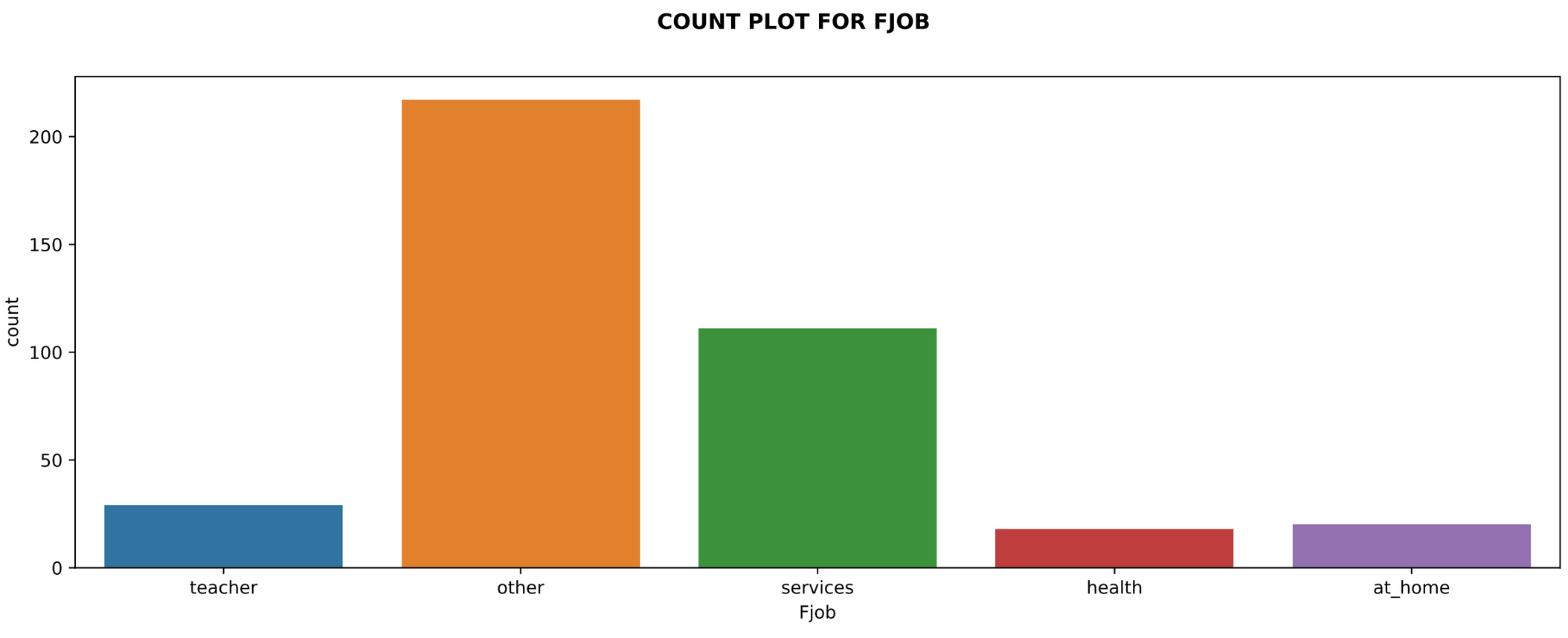

COUNT PLOT FOR FAMSIZE
count
0
50
100
150
200
250
GT3
LE3
famsize

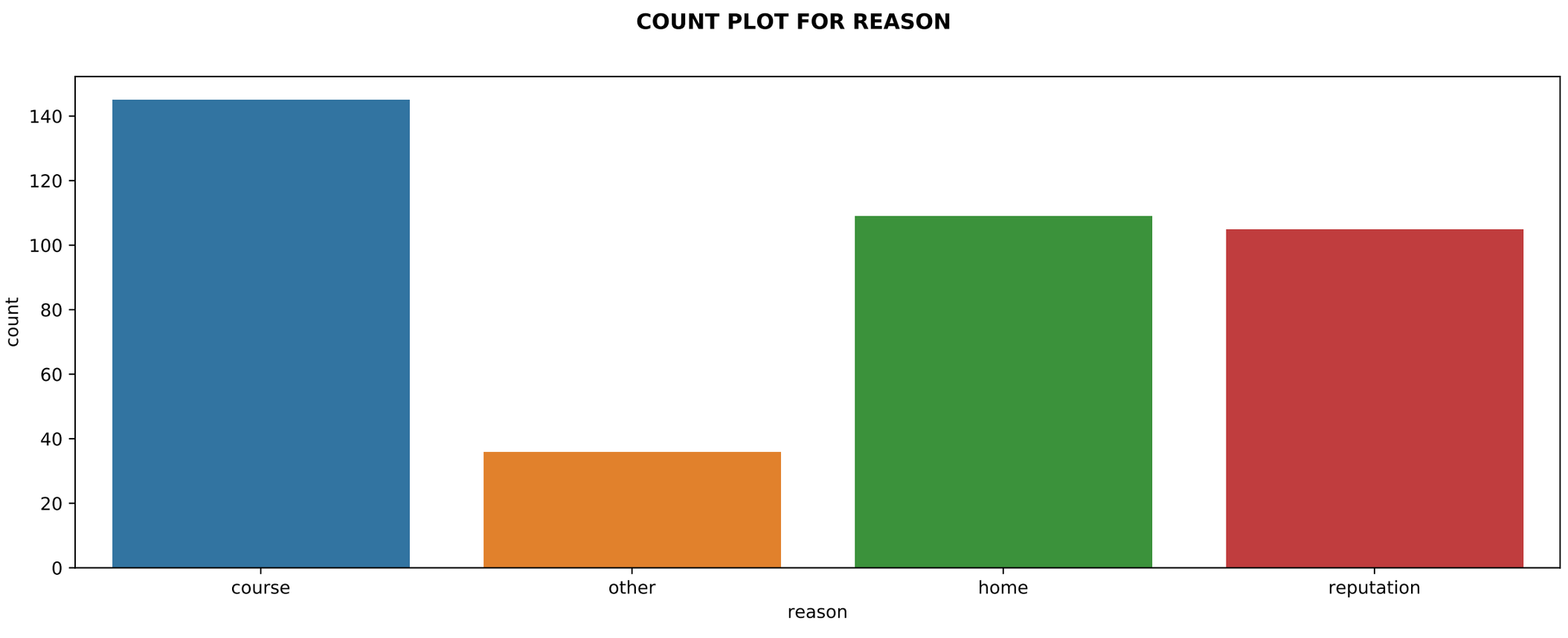

COUNT PLOT FOR PSTATUS
count
0
50
100
150
200
250
300
350
A
T
Pstatus

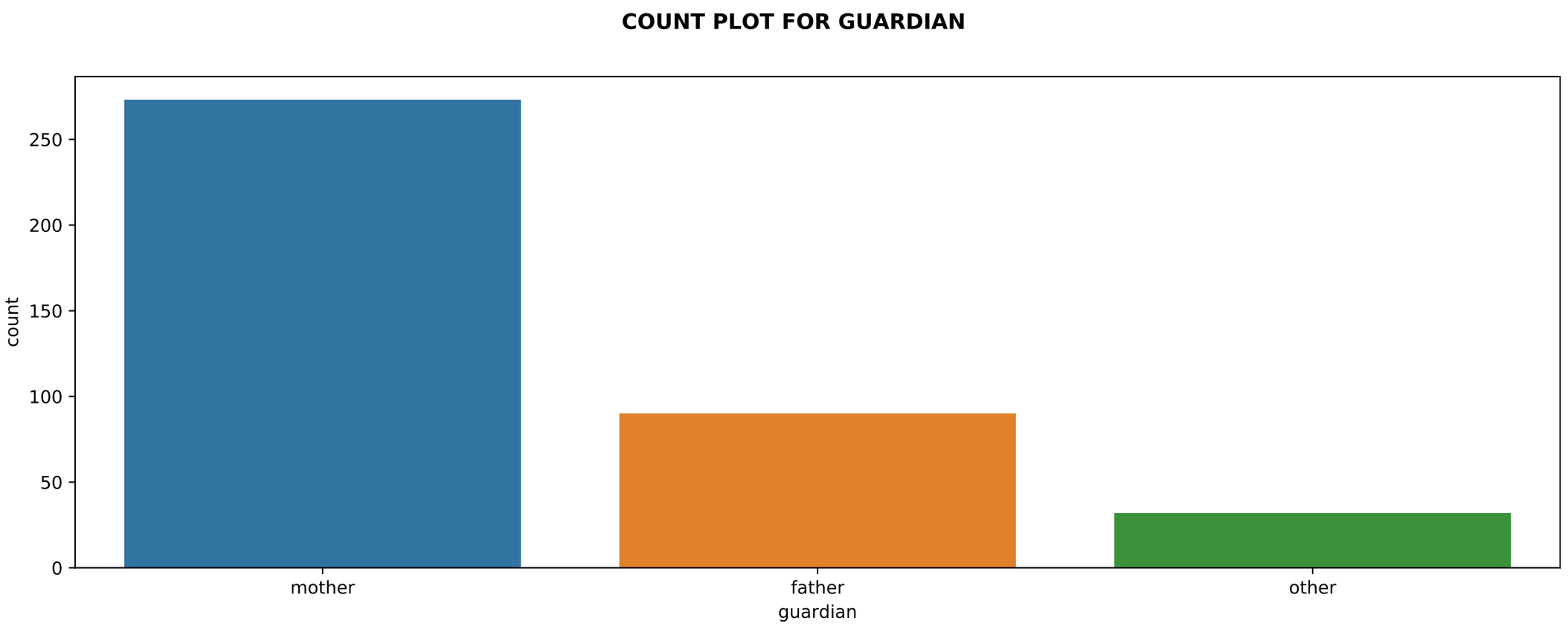

VARIOUS PLOTS FOR MEDU
Medu
0
Medu
0
Medu


VARIOUS PLOTS FOR FEDU
Fedu
0
Fedu
0
Fedu

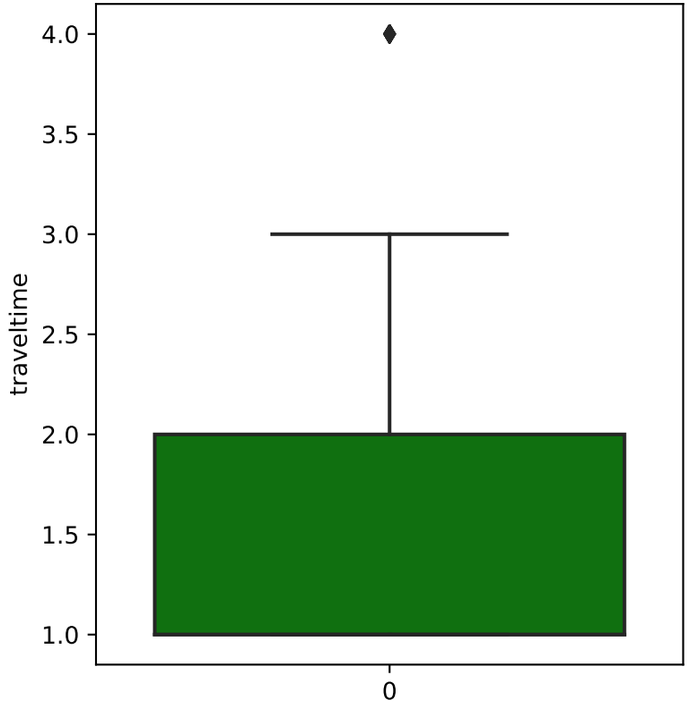

COUNT PLOT FOR MJOB
count
0
20
40
60
80
100
120
140
at_home
health
other
services
teacher
Mjob

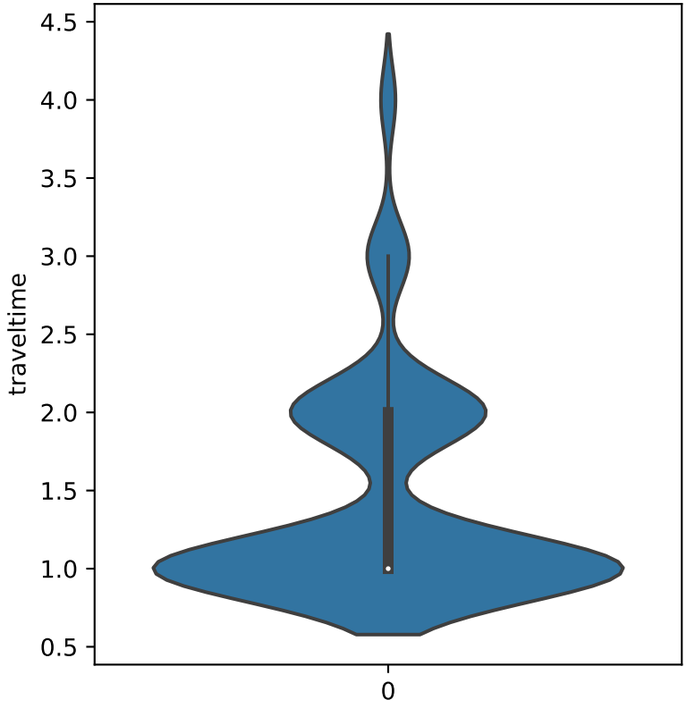

COUNT PLOT FOR FJOB
count
0
50
100
150
200
teacher
other
services
health
at_home
Fjob

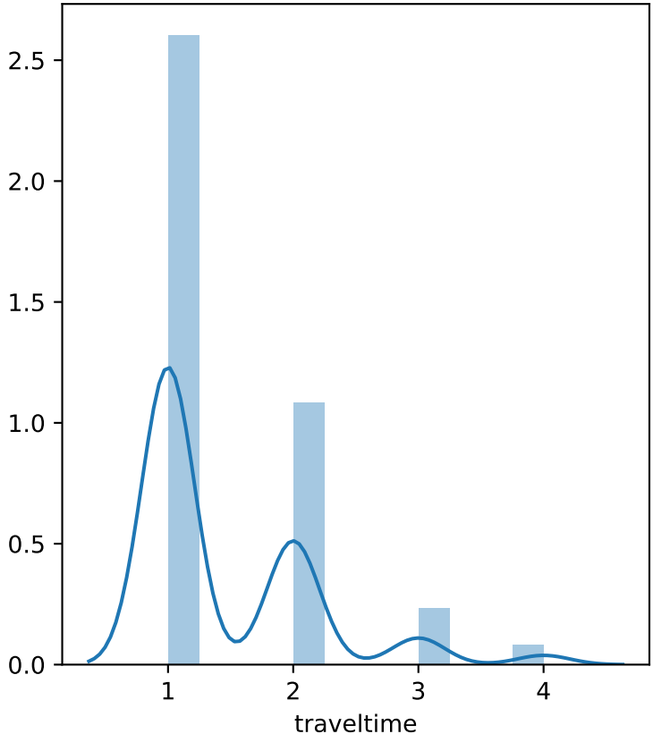

COUNT PLOT FOR REASON
count
0
20
40
60
80
100
120
140
course
other
home
reputation
reason

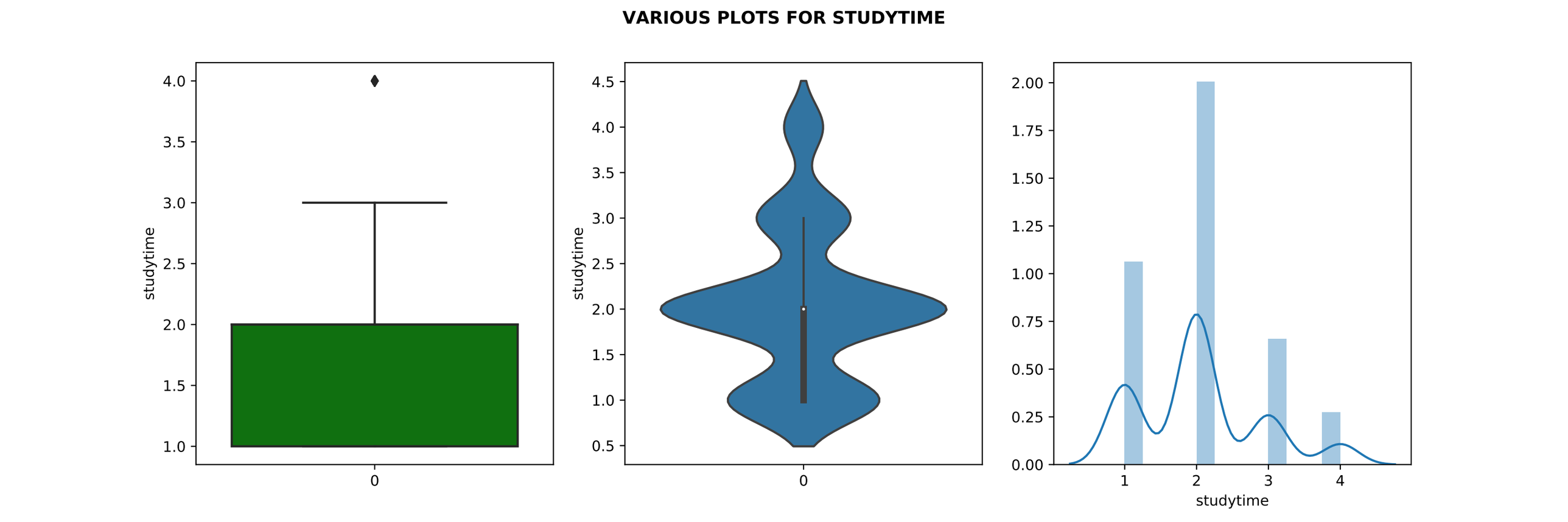

COUNT PLOT FOR GUARDIAN
count
0
50
100
150
200
250
mother
father
other
guardian

VARIOUS PLOTS FOR TRAVELTIME

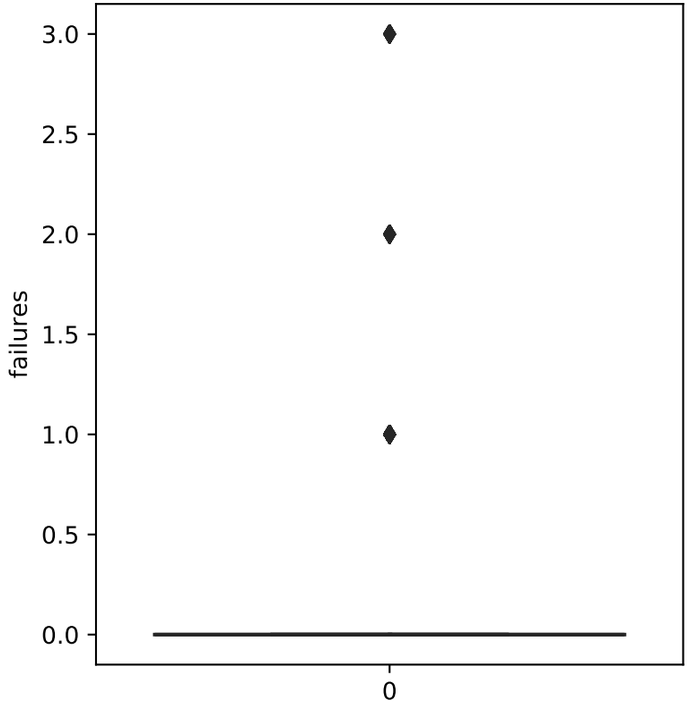


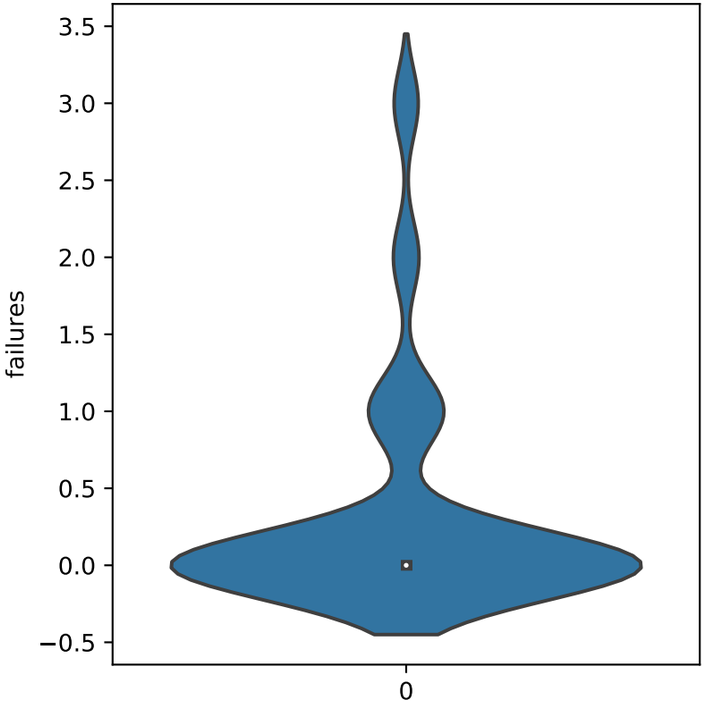


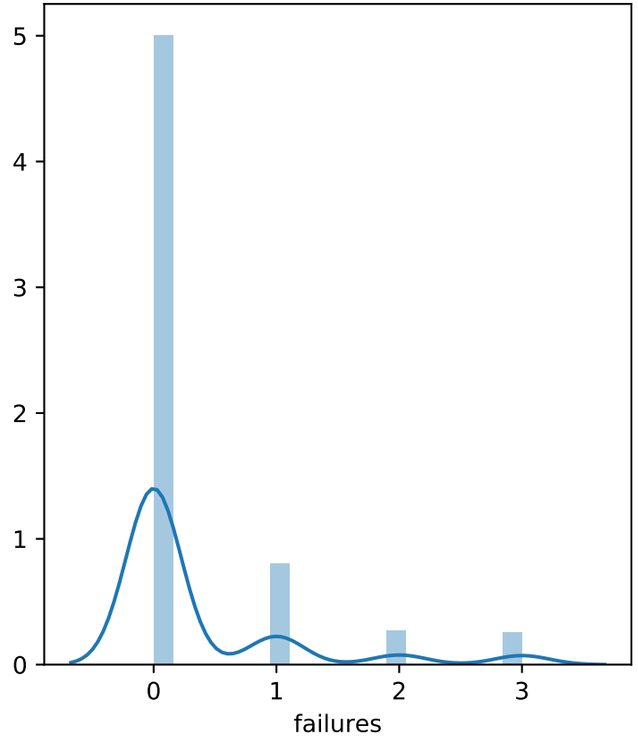


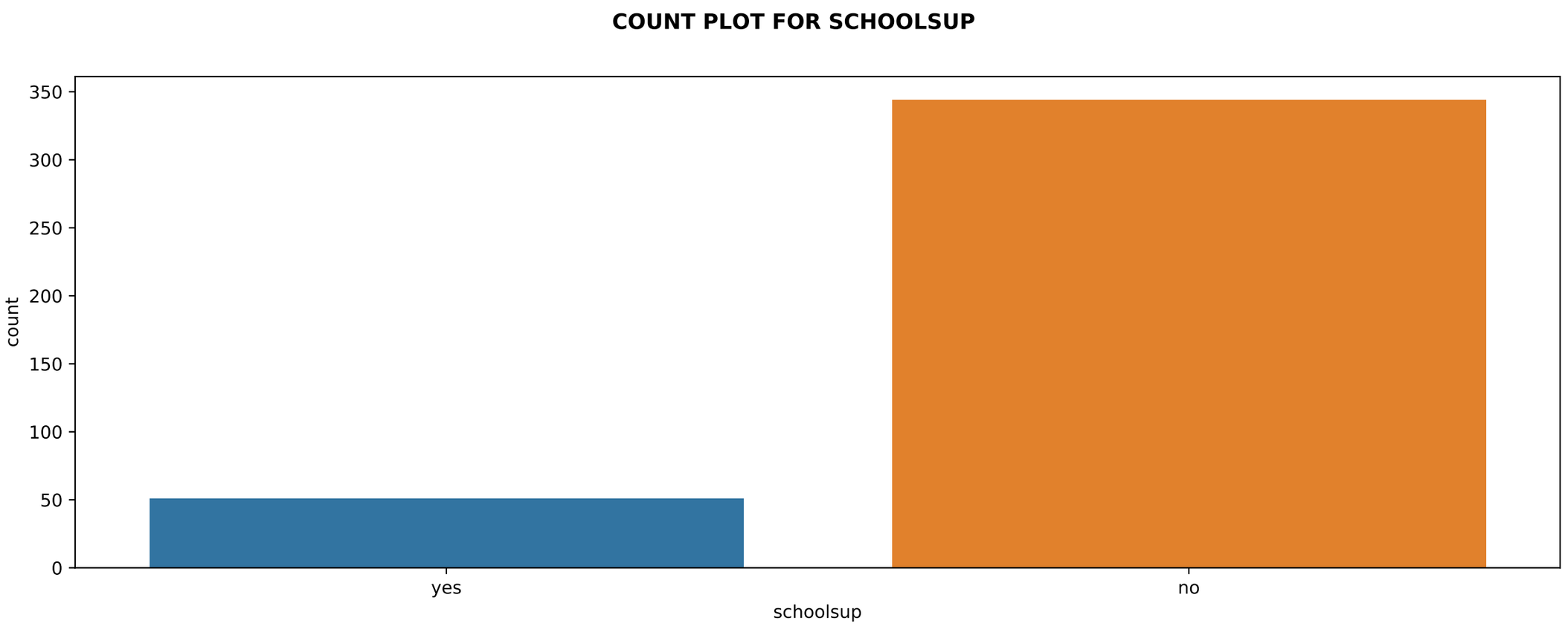

VARIOUS PLOTS FOR STUDYTIME
studytime
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
studytime
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
1
2
3
4
studytime

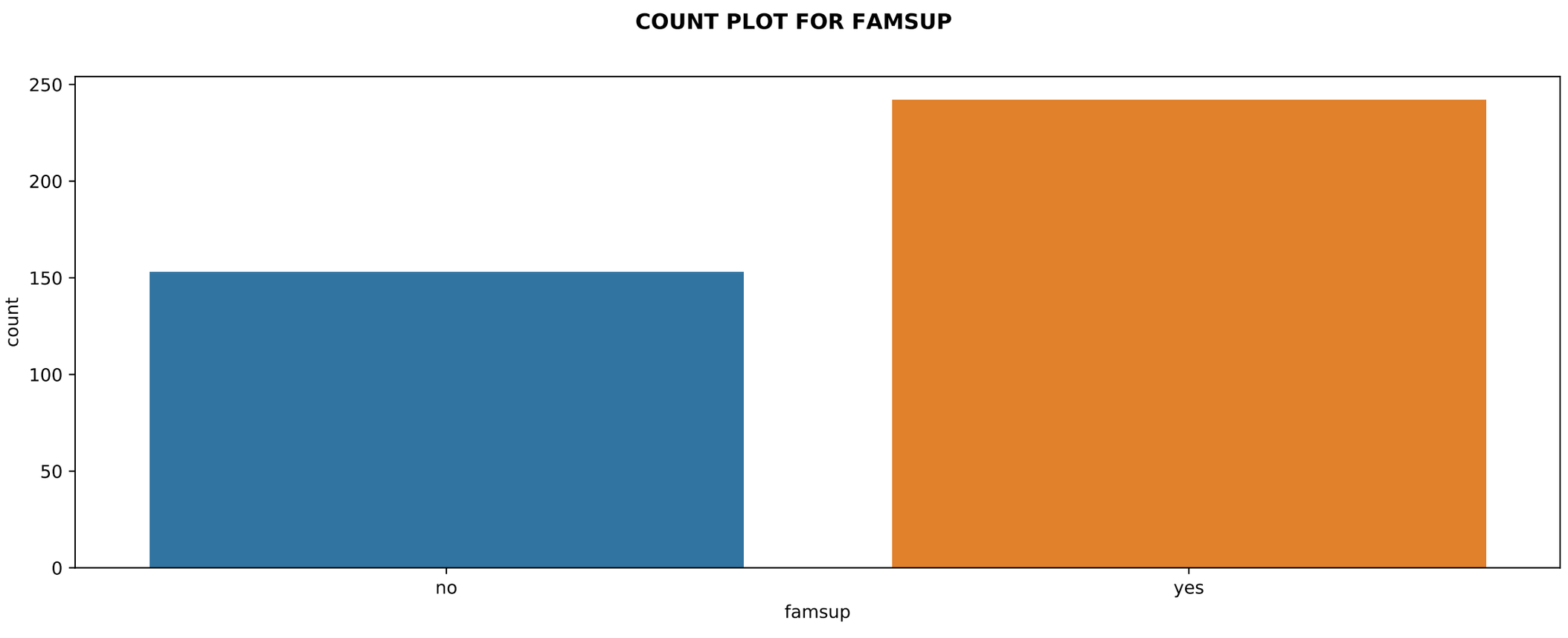
VARIOUS PLOTS FOR FAILURES

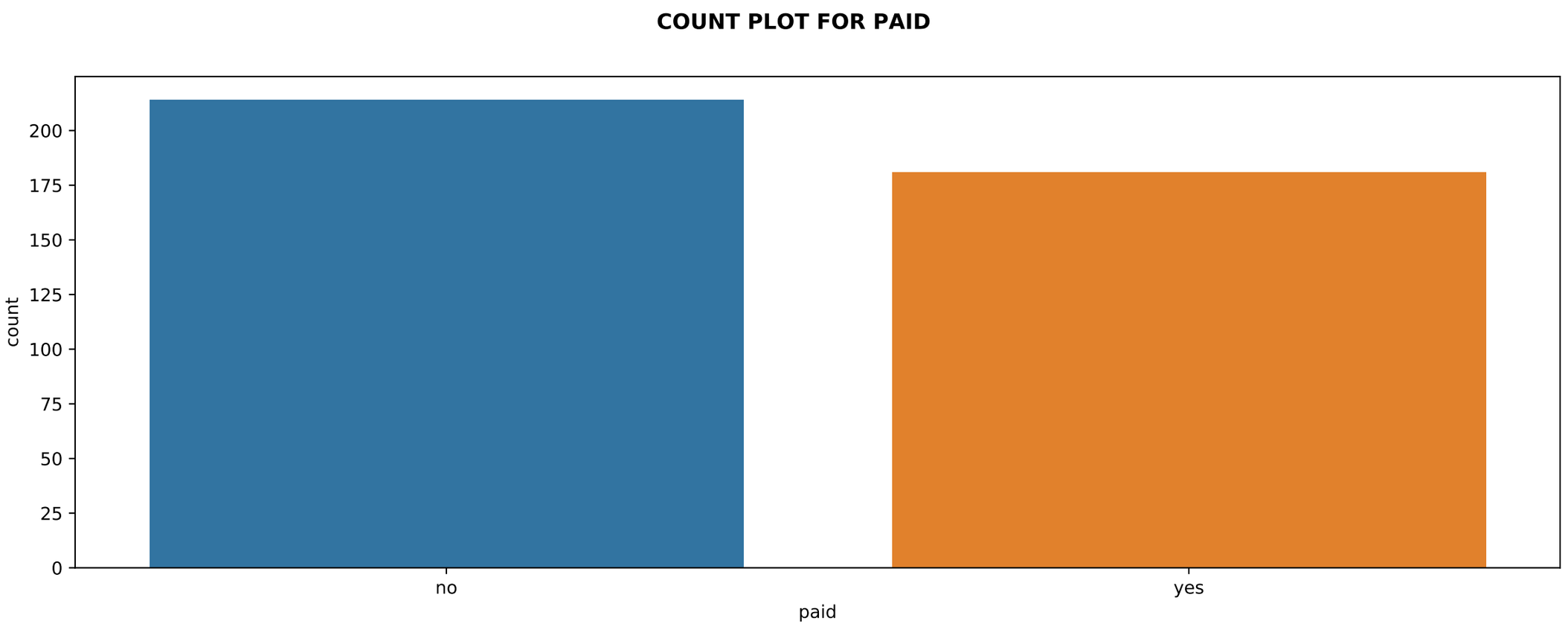


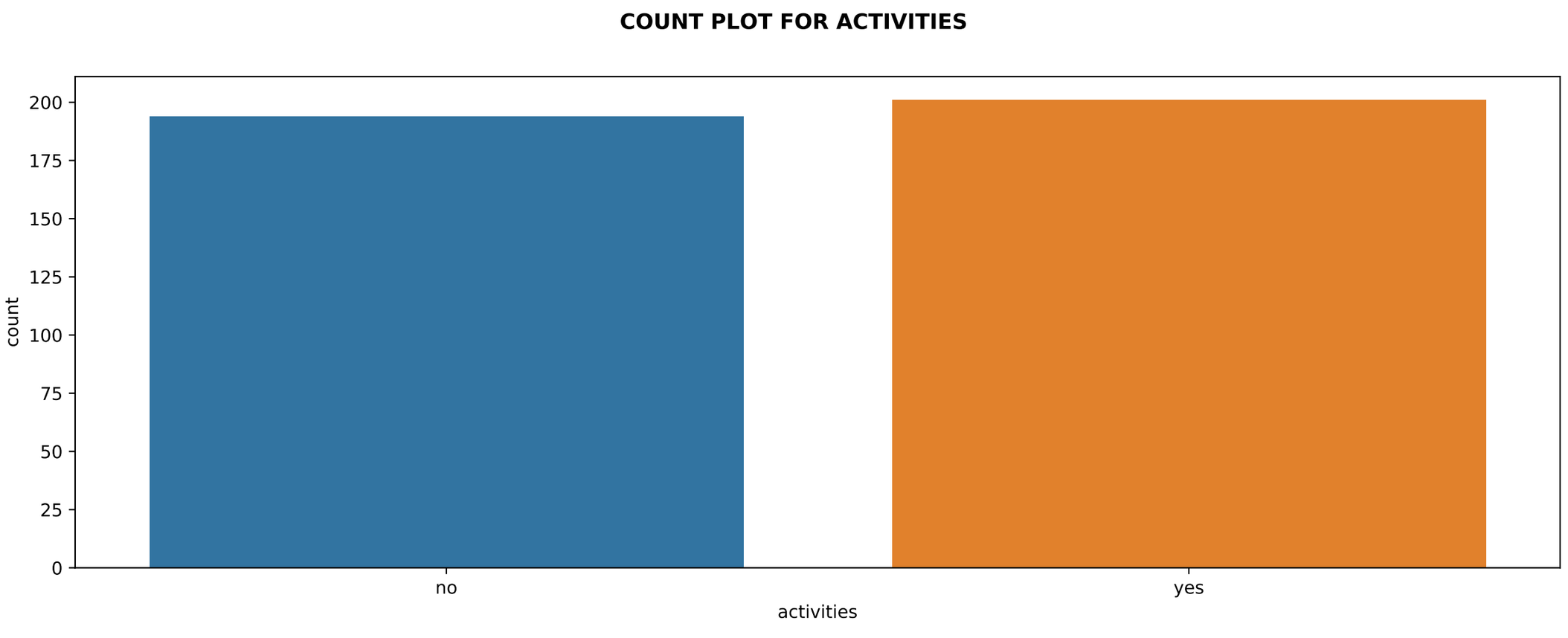


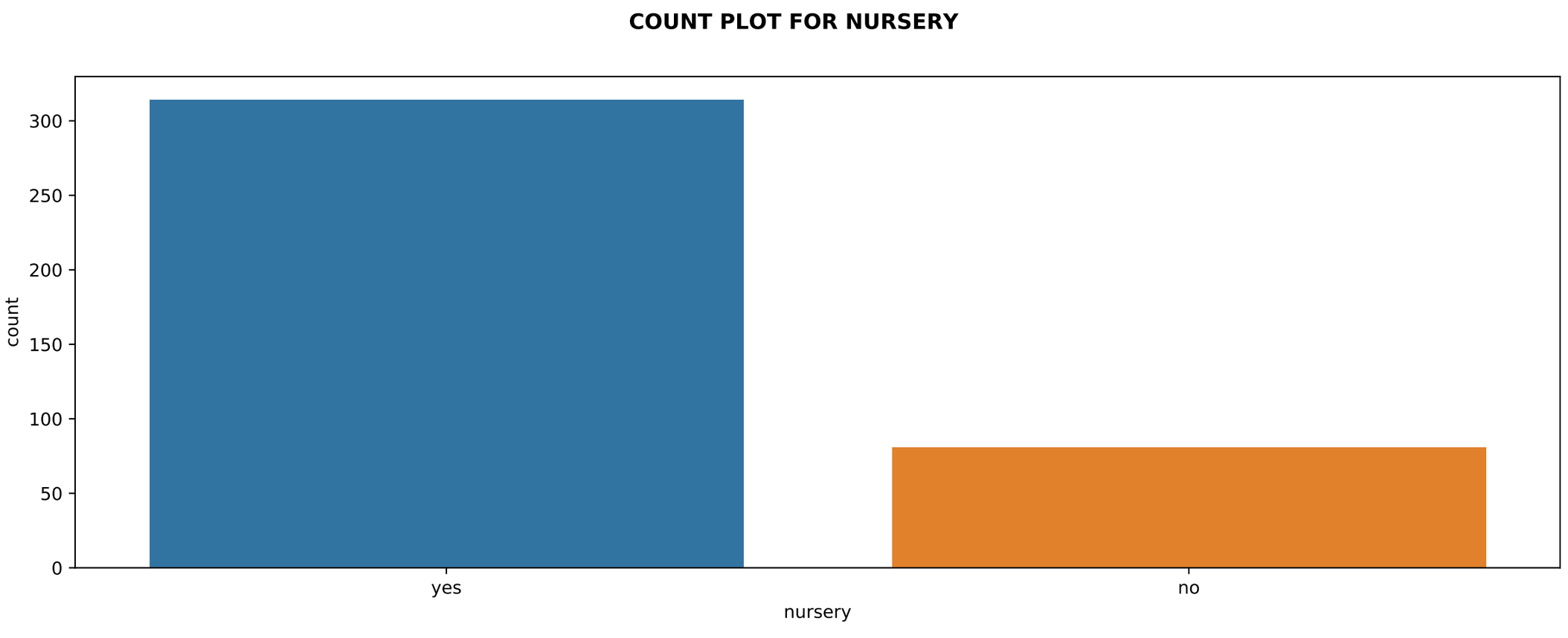


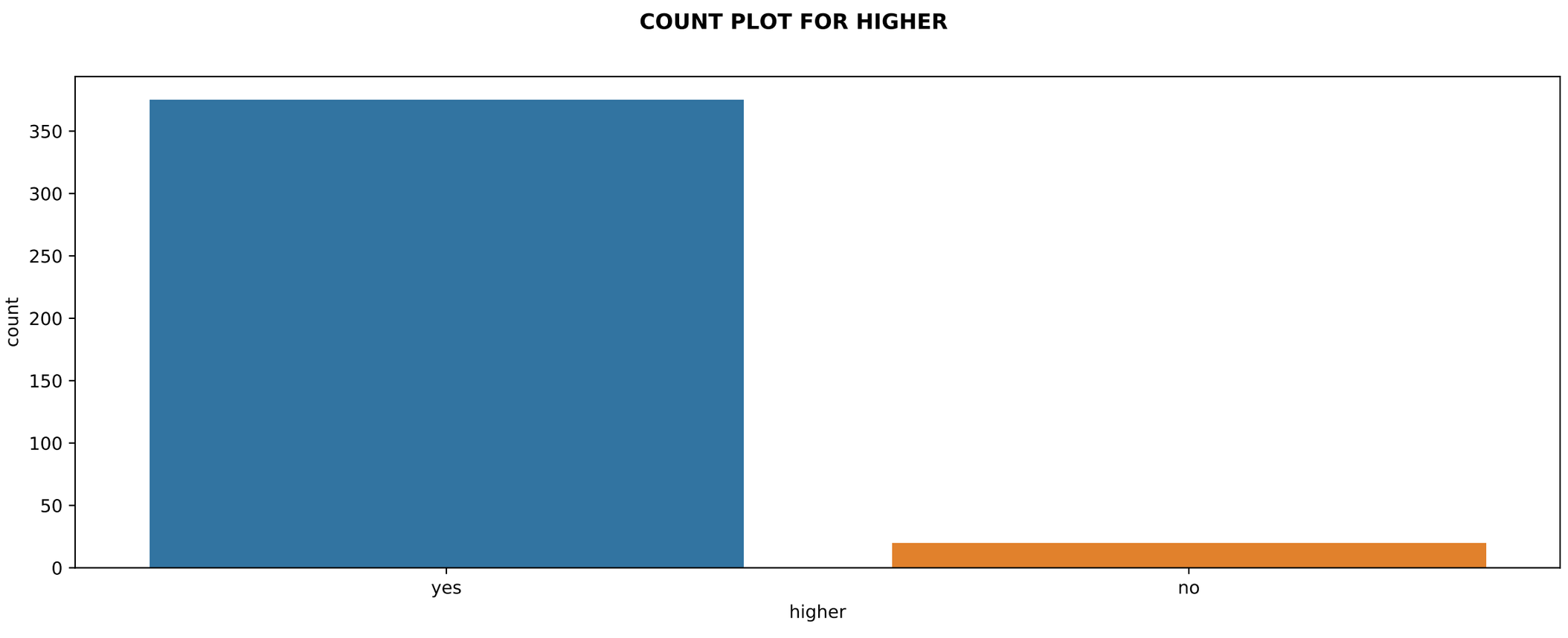

COUNT PLOT FOR SCHOOLSUP
350
300
250
200
150
100
50
0
count
yes
no
schoolsup

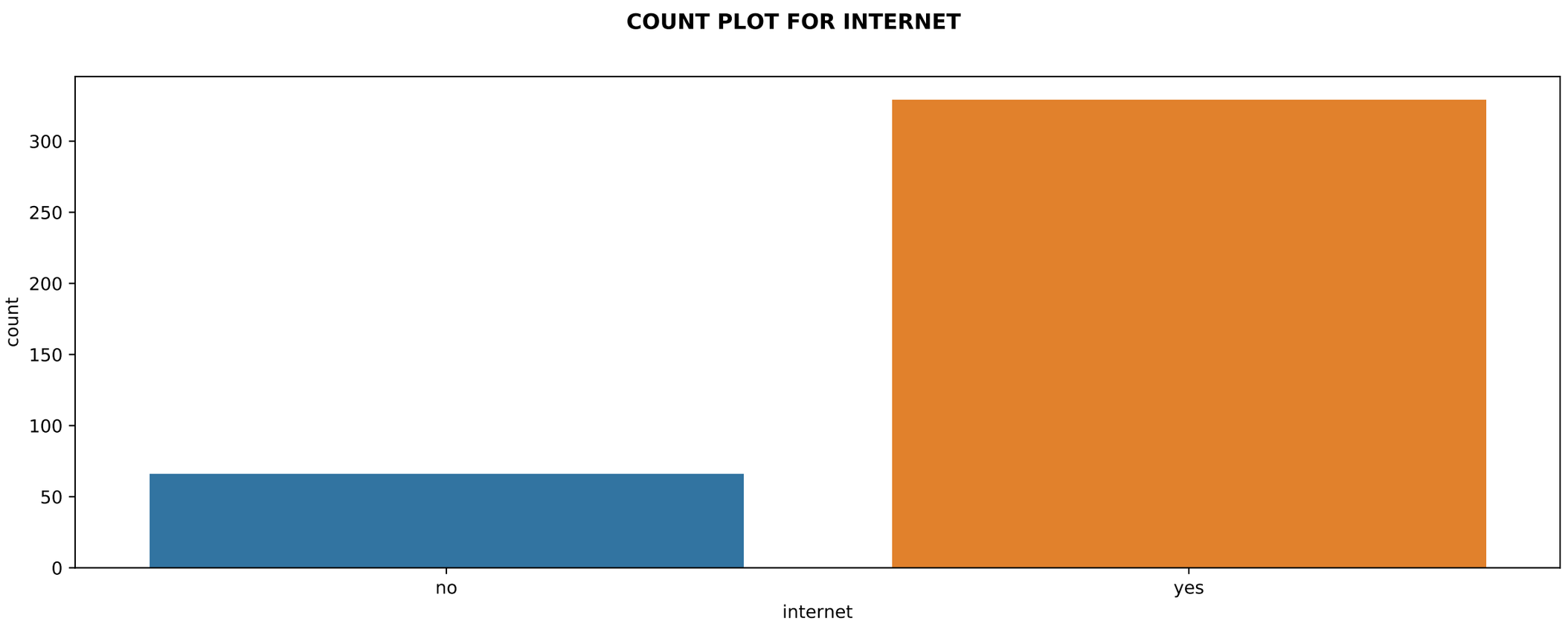

COUNT PLOT FOR FAMSUP
250
200
150
100
50
0
count
no
yes
famsup

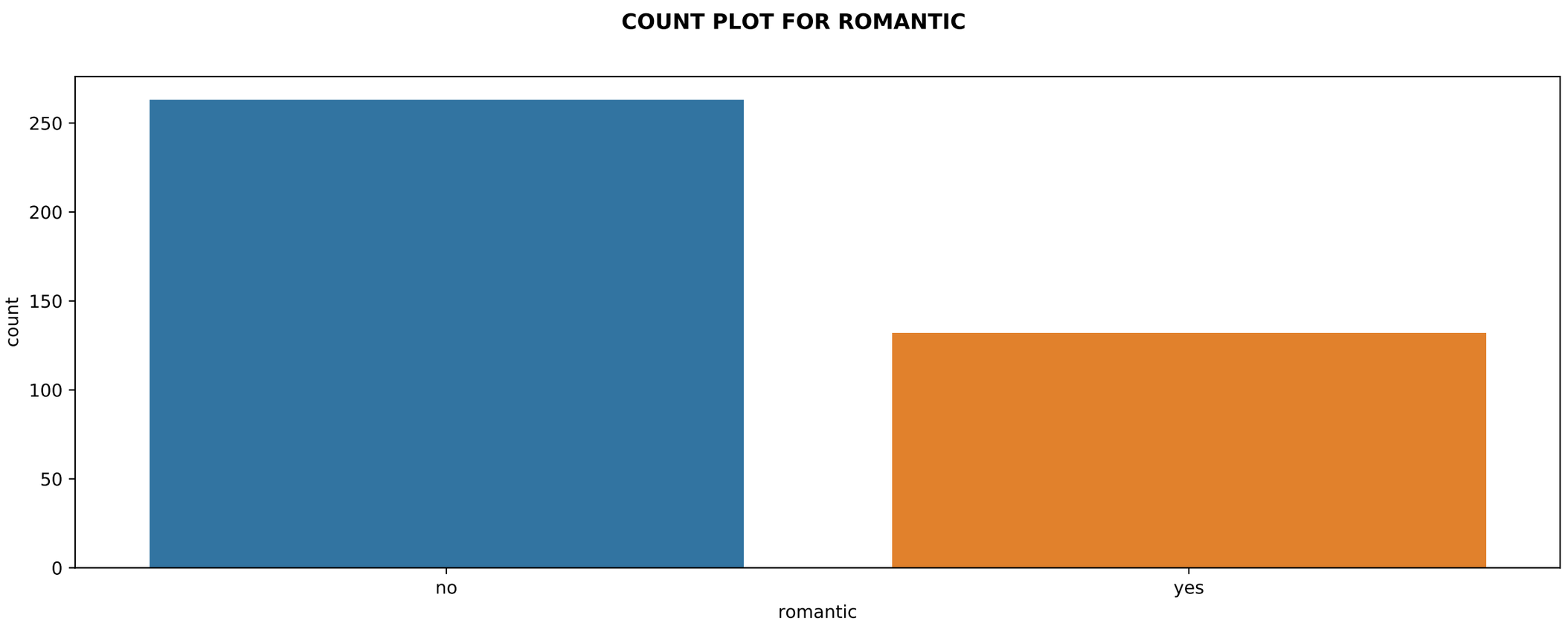

COUNT PLOT FOR PAID
count
0
25
50
75
100
125
150
175
200
no
yes
paid


COUNT PLOT FOR ACTIVITIES
200
175
150
125
100
75
50
25
0
count
no
yes
activities

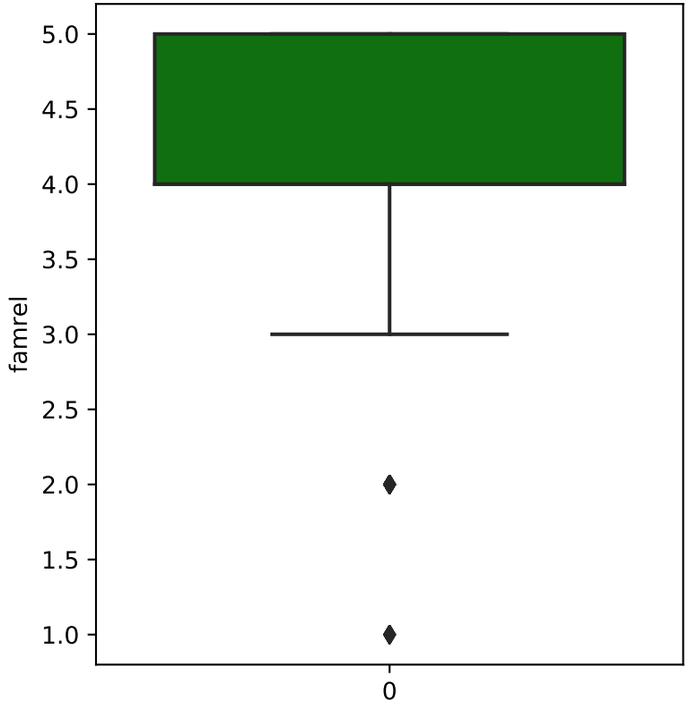

COUNT PLOT FOR NURSERY
count
0
50
100
150
200
250
300
yes
no
nursery

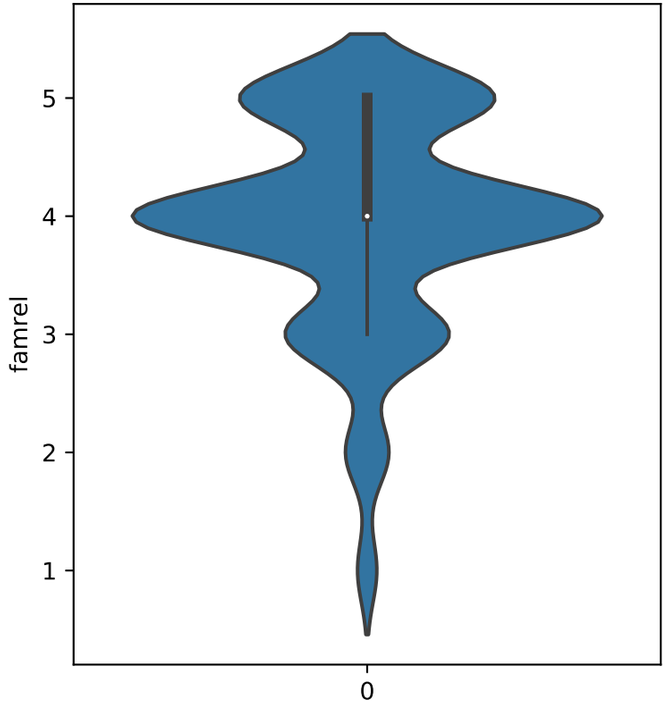

COUNT PLOT FOR HIGHER
count
0
50
100
150
200
250
300
350
yes
no
higher

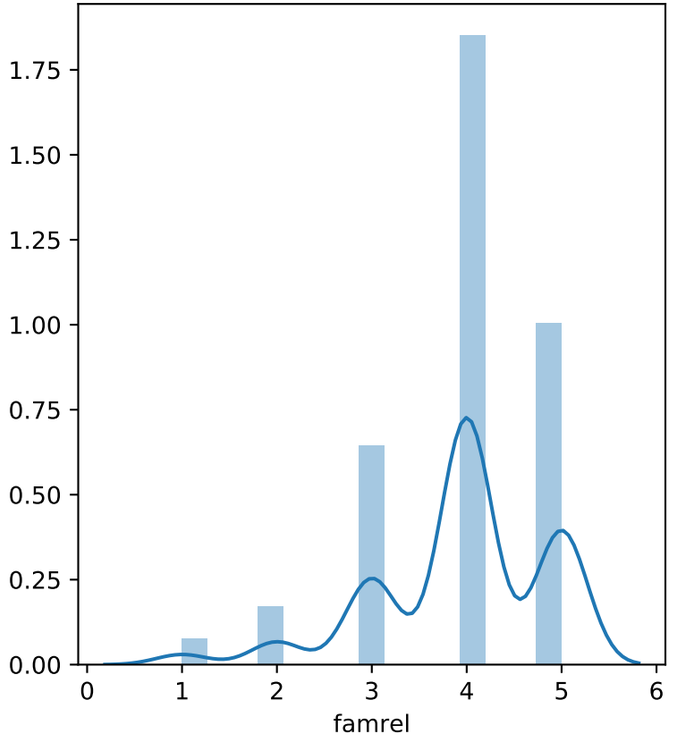

COUNT PLOT FOR INTERNET
count
0
50
100
150
200
250
300
no
yes
internet


COUNT PLOT FOR ROMANTIC
count
250
200
150
100
50
0
no
yes
romantic

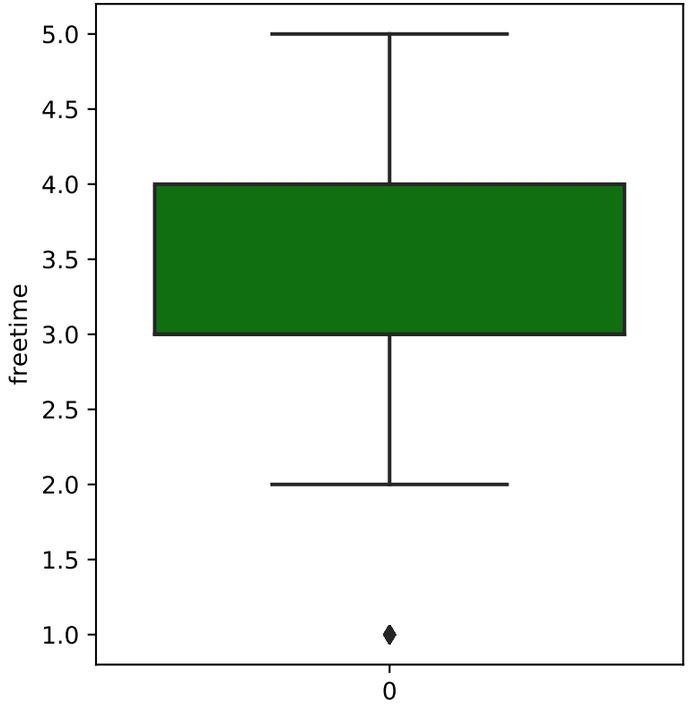
VARIOUS PLOTS FOR FAMREL

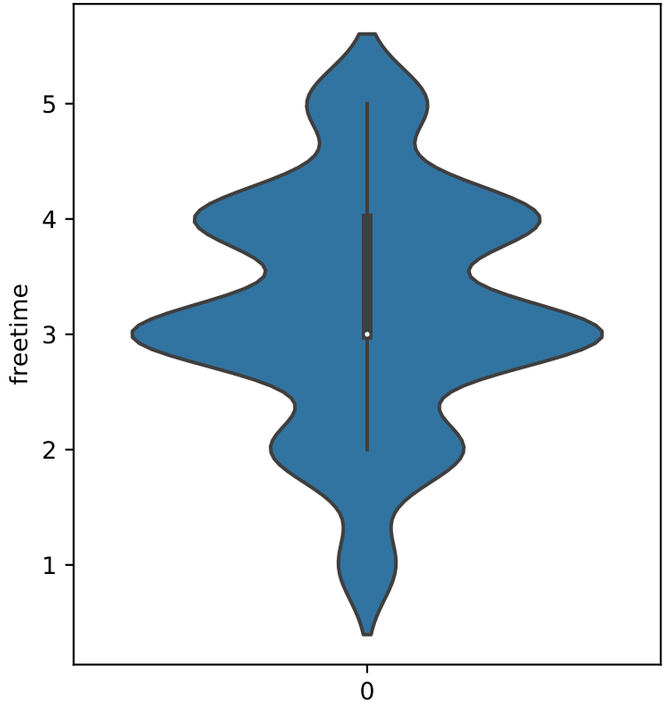


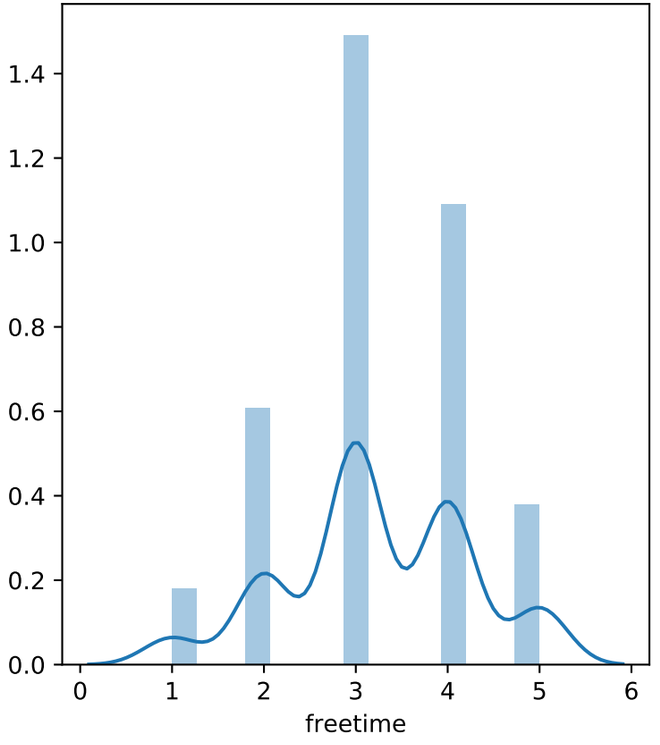


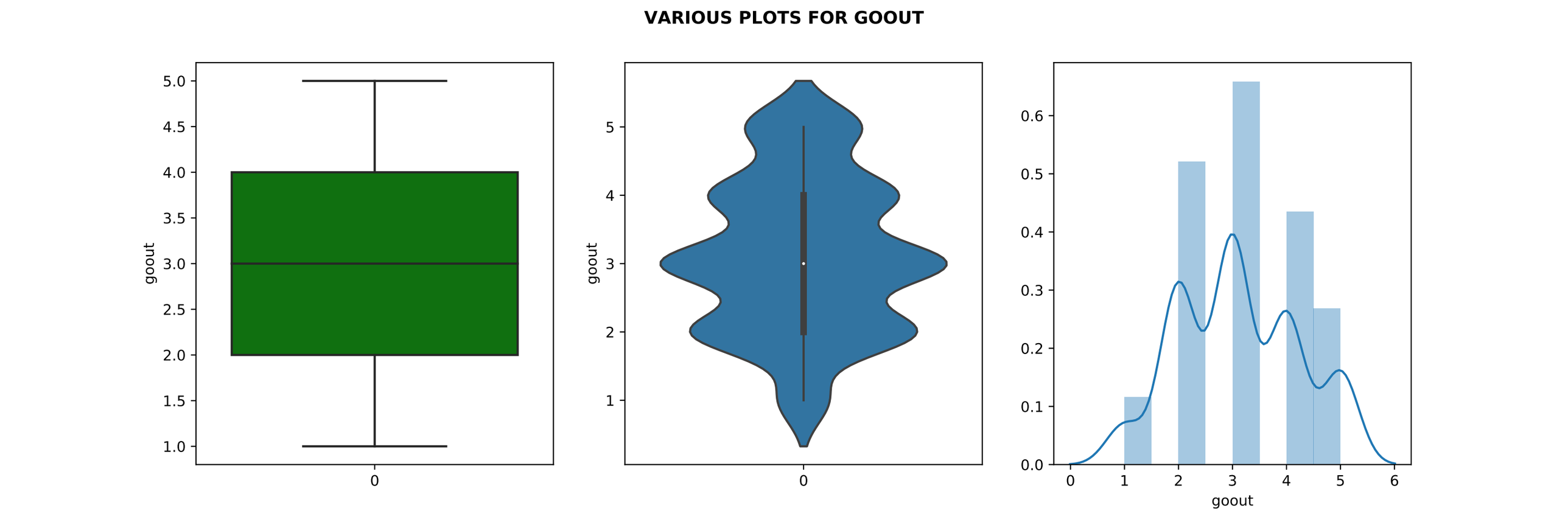


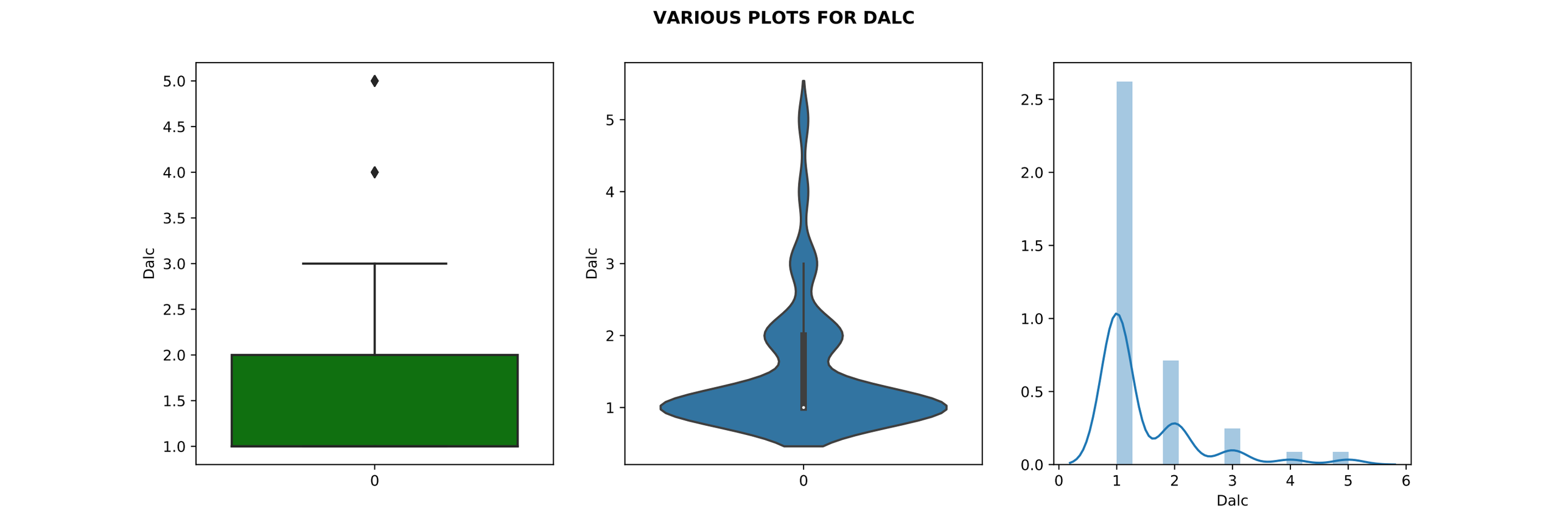
VARIOUS PLOTS FOR FREETIME

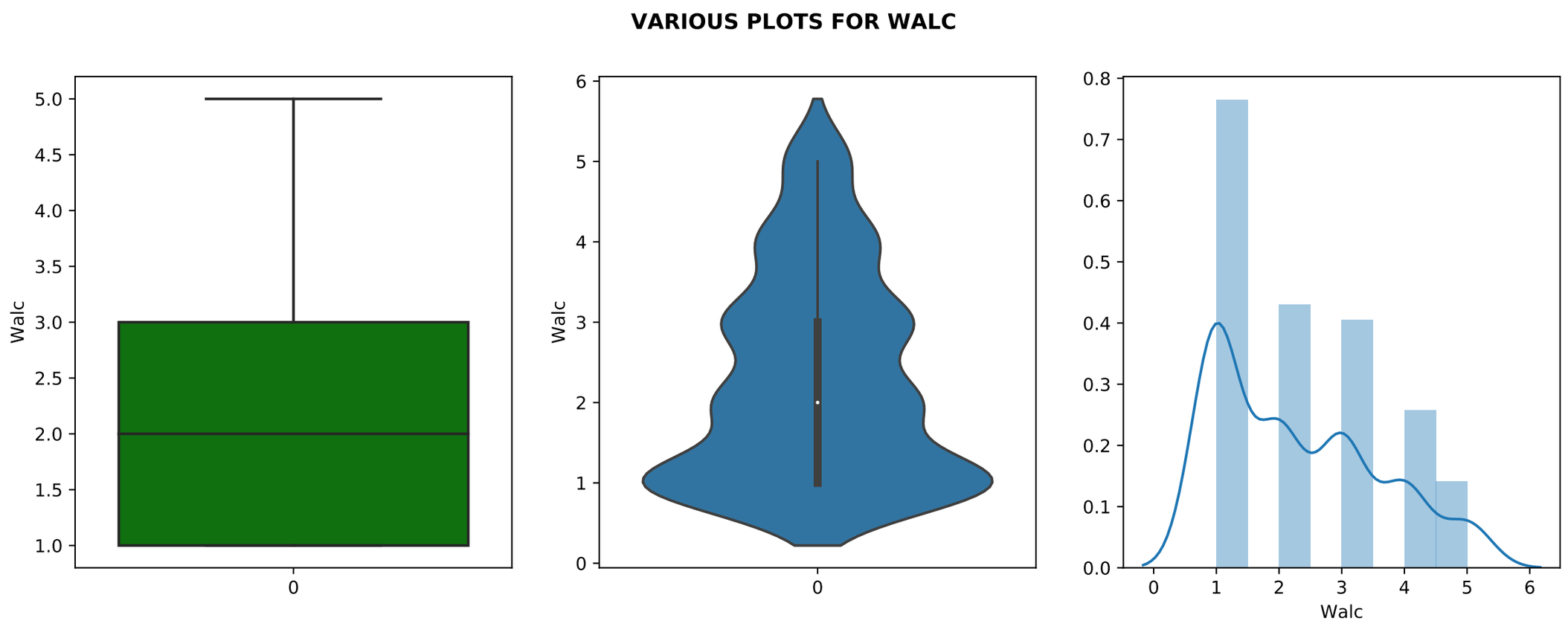




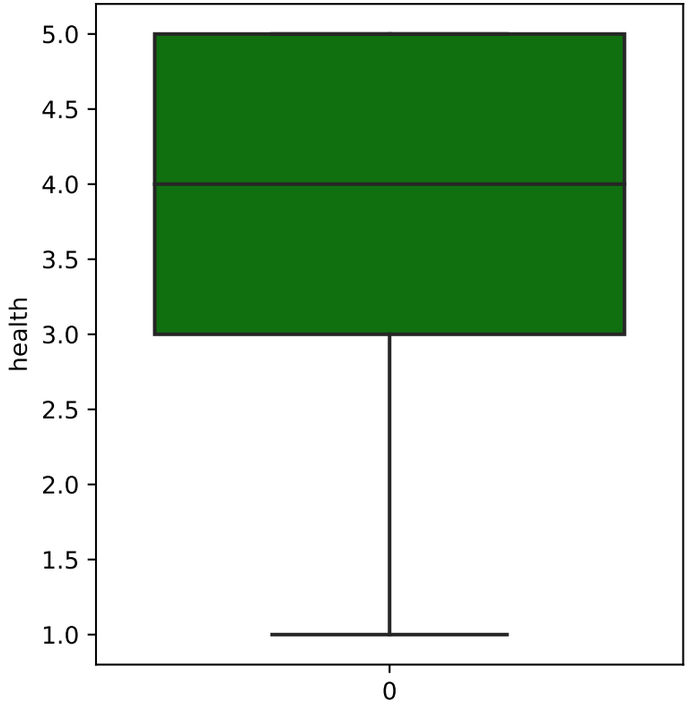


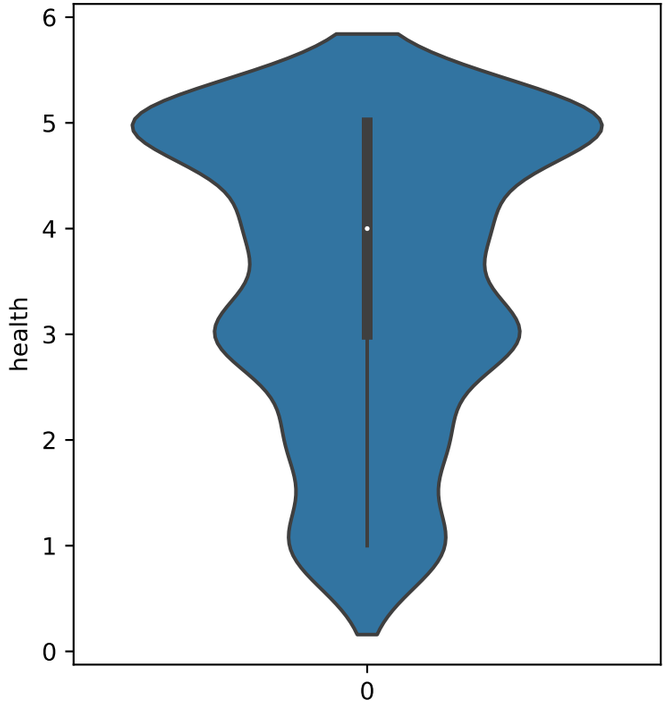

VARIOUS PLOTS FOR GOOUT
goout
0
goout
0
goout

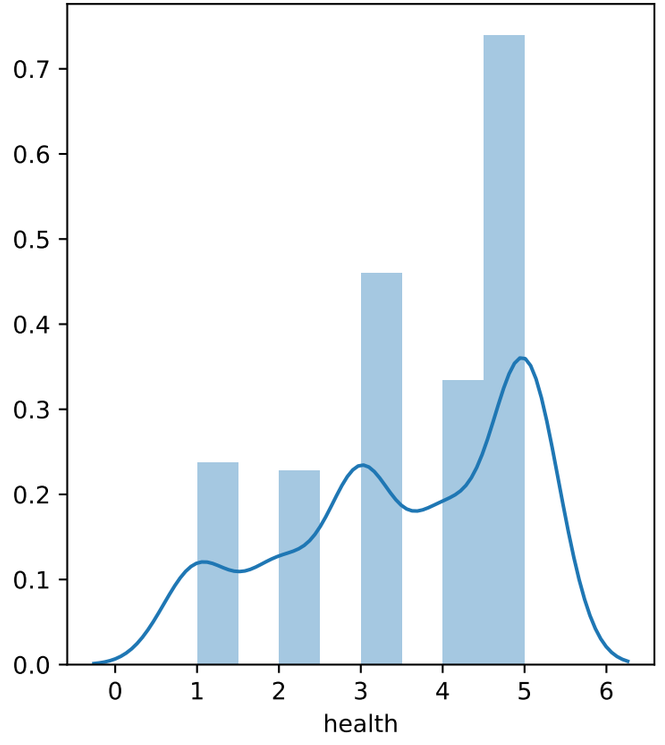

VARIOUS PLOTS FOR DALC
Dalc
0
Dalc
0
Dalc

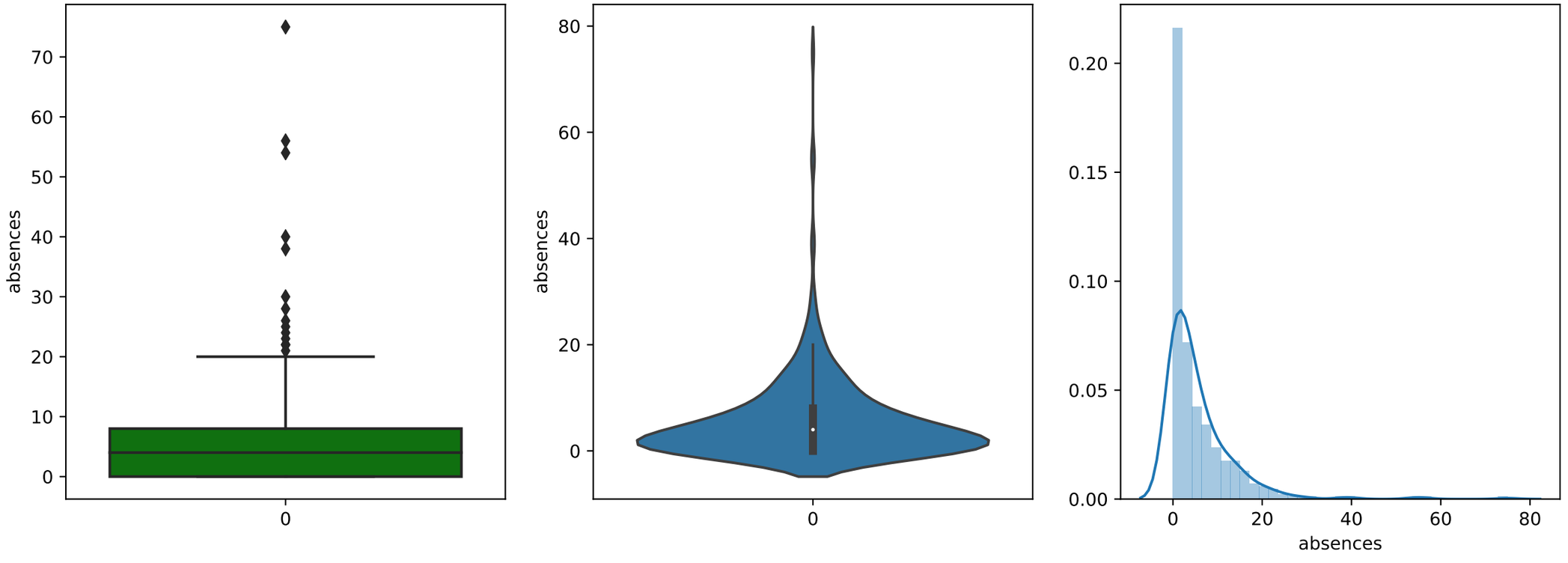

VARIOUS PLOTS FOR WALC
Walc
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
Walc
6
5
4
3
2
1
0
0
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
5
6
Walc

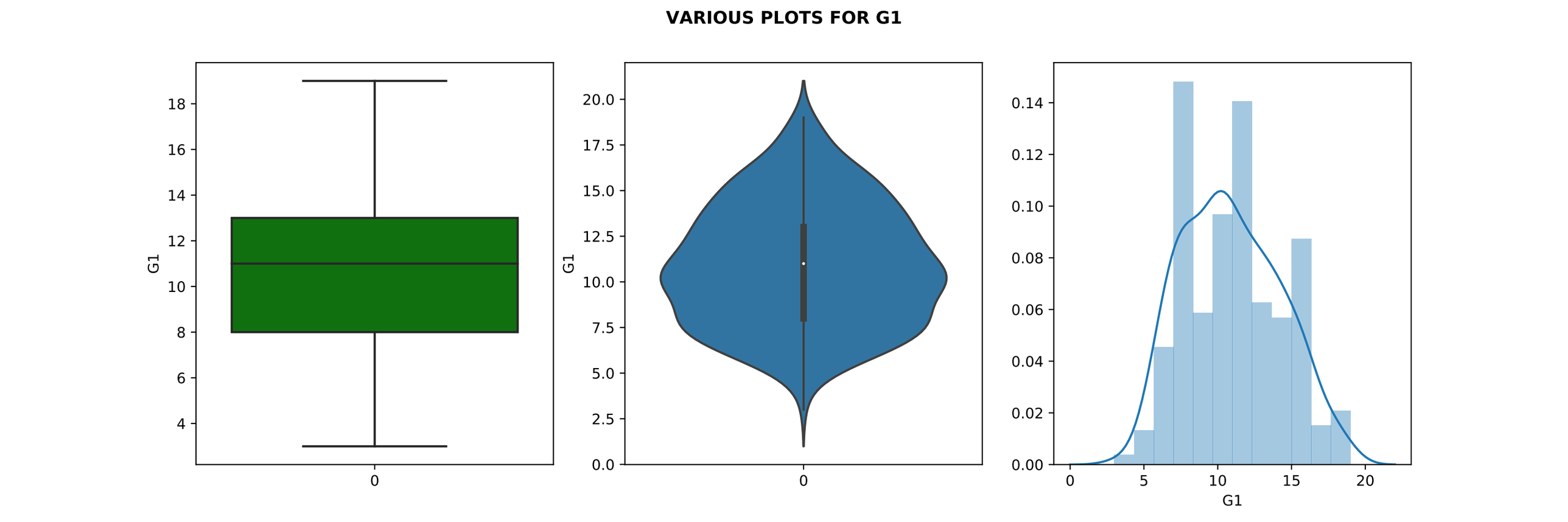
VARIOUS PLOTS FOR HEALTH

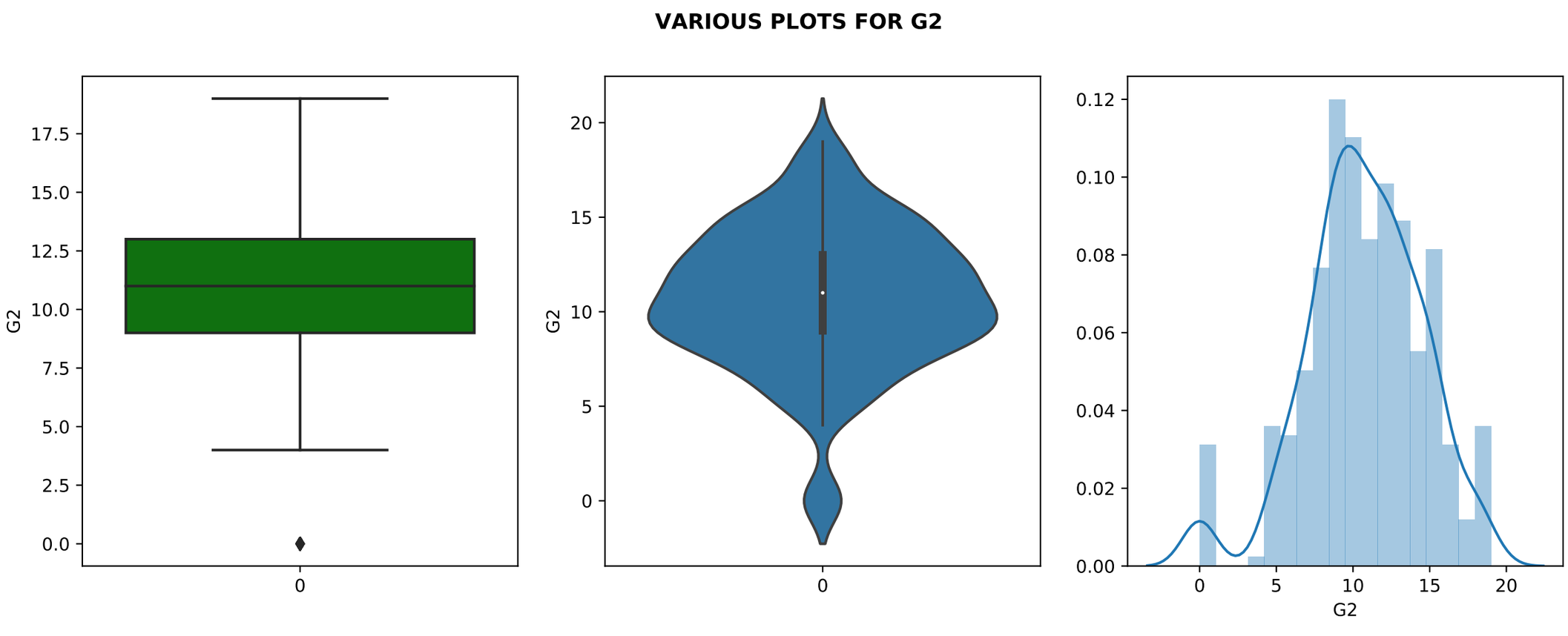







VARIOUS PLOTS FOR ABSENCES
absences
0
absences
0
absences

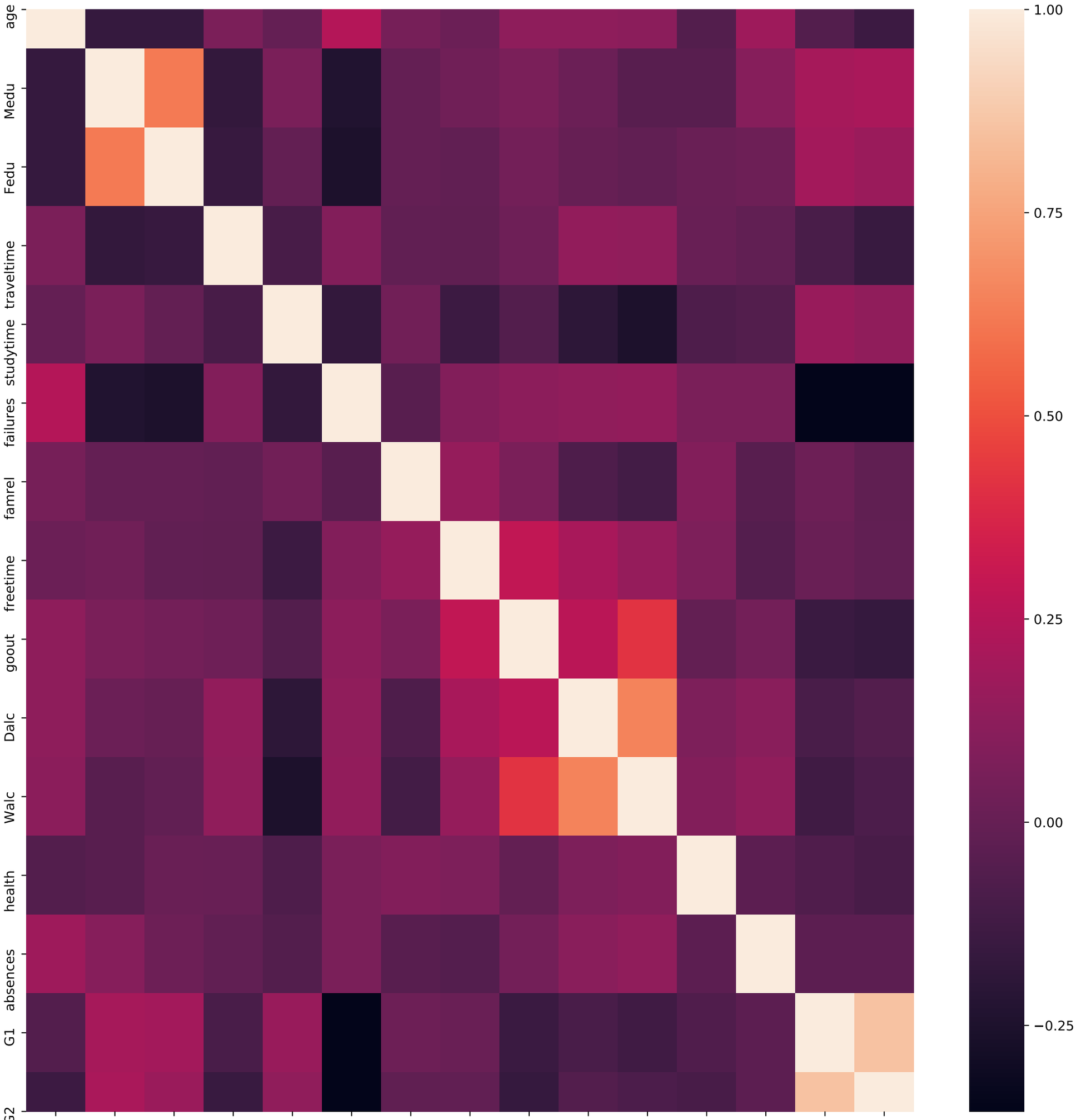

VARIOUS PLOTS FOR G1
G1
0
18
16
14
12
10
8
6
4
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
G1

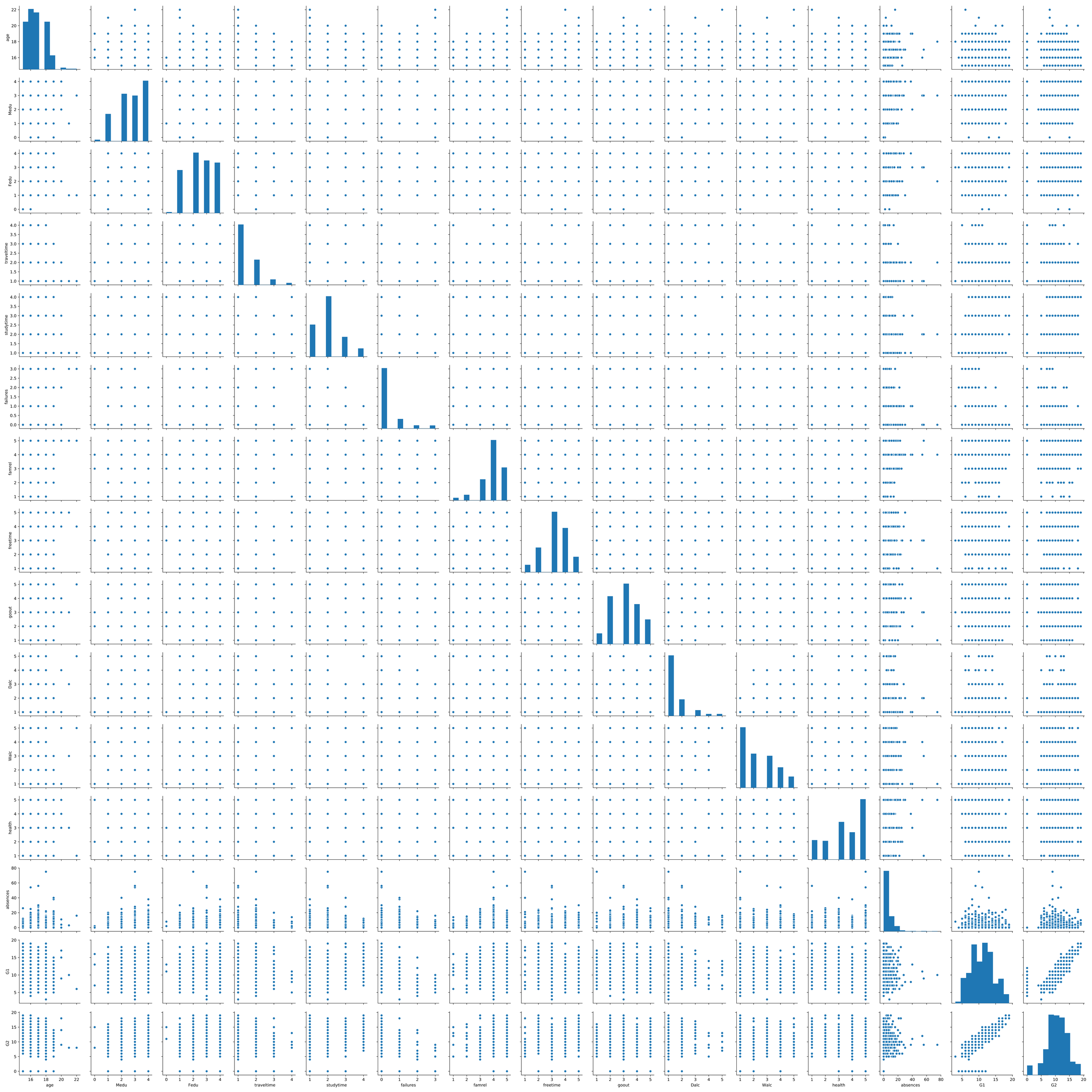

VARIOUS PLOTS FOR G2
G2
0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
20
15
10
5
0
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
5
10
15
20
G2

COORELATION MATRIX

age	Medu	Fedu	traveltime	studytime	failures	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2
1.0	-0.164	-0.163	0.071	-0.004	0.244	0.054	0.016	0.127	0.131	0.117	-0.062	0.175	-0.064	-0.143
-0.164	1.0	0.623	-0.172	0.065	-0.237	-0.004	0.031	0.064	0.02	-0.047	-0.047	0.1	0.205	0.216
-0.163	0.623	1.0	-0.158	-0.009	-0.25	-0.001	-0.013	0.043	0.002	-0.013	0.015	0.024	0.19	0.165
0.071	-0.172	-0.158	1.0	-0.101	0.092	-0.017	-0.017	0.029	0.138	0.134	0.008	-0.013	-0.093	-0.153
-0.004	0.065	-0.009	-0.101	1.0	-0.174	0.04	-0.143	-0.064	-0.196	-0.254	-0.076	-0.063	0.161	0.136
0.244	-0.237	-0.25	0.092	-0.174	1.0	-0.044	0.092	0.125	0.136	0.142	0.066	0.064	-0.355	-0.356
0.054	-0.004	-0.001	-0.017	0.04	-0.044	1.0	0.151	0.065	-0.078	-0.113	0.094	-0.044	0.022	-0.018
0.016	0.031	-0.013	-0.017	-0.143	0.092	0.151	1.0	0.285	0.209	0.148	0.076	-0.058	0.013	-0.014
0.127	0.064	0.043	0.029	-0.064	0.125	0.065	0.285	1.0	0.267	0.42	-0.01	0.044	-0.149	-0.162
0.131	0.02	0.002	0.138	-0.196	0.136	-0.078	0.209	0.267	1.0	0.648	0.077	0.112	-0.094	-0.064
0.117	-0.047	-0.013	0.134	-0.254	0.142	-0.113	0.148	0.42	0.648	1.0	0.092	0.136	-0.126	-0.085
-0.062	-0.047	0.015	0.008	-0.076	0.066	0.094	0.076	-0.01	0.077	0.092	1.0	-0.03	-0.073	-0.098
0.175	0.1	0.024	-0.013	-0.063	0.064	-0.044	-0.058	0.044	0.112	0.136	-0.03	1.0	-0.031	-0.032
-0.064	0.205	0.19	-0.093	0.161	-0.355	0.022	0.013	-0.149	-0.094	-0.126	-0.073	-0.031	1.0	0.852
-0.143	0.216	0.165	-0.153	0.136	-0.356	-0.018	-0.014	-0.162	-0.064	-0.085	-0.098	-0.032	0.852	1.0

CORRELATION HEATMAP




age
Medu
Fedu
traveltime
studytime
failures
famrel
freetime
goout
Dalc
Walc
health
absences
G1
G2